

STIC-Biotech/ChemLib

56371

From: Nickol, Gary  
Sent: Saturday, December 08, 2001 11:36 AM  
To: STIC-Biotech/ChemLib  
Subject: 09/680121

Please search and Interference search the following:

- 1) DNA of SEQ ID NO:1
- 2) DNA encoding SEQ ID NO:2
- 3) Amino acids of SEQ ID NO:2

Thanks,

Gary Nickol Ph.D.  
AU:1642, Room 8D09, Mailbox 8E12  
703-305-7143

Point of Contact:  
Toby Port  
Technical Info. Specialist  
CM1 4E01 TEL: 308-3534  
12C14

RECEIVED  
DEC 10 2001  
STIC

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 12/12  
Date Completed: 12/14  
Searcher Prep/Review: 10  
Clerical: \_\_\_\_\_  
Online time: 10

TYPE OF SEARCH:  
NA Sequences: 2  
AA Sequences: 1  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST(where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: eg  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

THIS PAGE BLANK (USPTO)

Point of Contact:  
Toby Port  
Technical Info. Specialist  
CALL 1E01 TEL: 308-3234

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2001, 10:54:04 ; Search time 5414.84 Seconds  
(without alignments)  
11854.556 Million cell updates/sec

Title: US-09-680-121-1

Perfect score: 3891

Sequence: 1 cctcttcgccctccctcgtt.....ccccaaaaaaaaaaaaa 3891

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pal:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_ov:\*  
21: em\_or:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htgo\_rod:\*  
33: em\_htgo\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3891	100.0	3891	6	ARI46667	ARI46667 Sequence
2	3830.2	98.4	3910	9	AB037763	AB037763 Homo sapi
3	2720.4	69.9	17117	2	AC009266	AC009266 Homo sapi
4	2720.4	69.9	190550	2	AP001797	AP001797 Homo sapi
5	2720.4	69.9	192022	2	AC087507	AC087507 Homo sapi
6	2709.4	69.6	159135	2	AC091039	AC091039 Homo sapi
7	1651	42.4	1671	9	AF299075	AF299075 Homo sapi
8	1571.8	40.4	170218	2	AC068125	AC068125 Homo sapi
9	1305.4	33.5	3992	10	MMU10355	MMU10355 Mus muscu
10	1124.6	28.9	2060	10	NMU14398	NMU14398 Rattus norv
11	1054.2	27.1	1576	10	RATPST	L38247 Rattus norv
12	363.6	9.3	2950	9	BC004291	BC004291 Homo sapi
13	362	9.3	2315	9	AK027340	AK027340 Homo sapi
14	346.2	8.9	2426	10	AF000423	AF000423 Rattus no
15	346	8.9	1293	10	AB026808	AB026808 Mus muscu
16	344.4	8.9	1293	10	AF375465	AF375465 Rattus no
17	281.8	7.2	286	11	G24341	G24341 human STR W
18	207.8	5.3	148316	2	AL356483	AL356483 Homo sapi
19	207.8	5.3	163983	2	AL139128	AL139128 Homo sapi
20	164.4	4.2	1163	10	AB026806	AB026806 Mus muscu
21	163.2	4.2	1876	10	MUSSYNB	D37793 Mouse mRNA
22	163.2	4.2	1876	22	E11219	E11219 Mouse cDNA
23	163.2	4.2	2544	3	AB044144	AB044144 Halocynthia
24	163.2	4.2	2795	10	AF257304	AF257304 Mus muscu
25	163	4.2	1417	3	ACU03125	U03125 Aplysia cal
26	159.4	4.1	1277	3	LPSYNTGM	X72386 L. pealei mr
27	157.8	4.1	1900	3	D63797	D63797 Loligo peal
28	157	4.0	2551	5	S64957	S64957 synaptotagm
29	156.4	4.0	1440	10	RNU20106	RNU20106 Rattus norv
30	156.2	4.0	1212	10	AB026804	AB026804 Mus muscu
31	156.2	4.0	1212	10	AF336854	AF336854 Rattus no
32	156.2	4.0	1563	10	AF336856	AF336856 Rattus no
33	156.2	4.0	1572	10	AF336857	AF336857 Rattus no
34	156.2	4.0	1704	10	AF336858	AF336858 Rattus no
35	156.2	4.0	1836	10	AF336859	AF336859 Rattus no
36	156.2	4.0	1932	10	AF336860	AF336860 Rattus no
37	156.2	4.0	1959	10	AF336861	AF336861 Rattus no
38	156.2	4.0	2064	10	AF336865	AF336865 Rattus no
39	156.2	4.0	2187	10	AF336862	AF336862 Rattus no
40	155	4.0	870	10	RNU20108	RNU20108 Rattus norv
41	153.6	3.9	1620	10	RNRNASYNV	X84884 R. norvegicu
42	153.6	3.9	1639	10	RNU26402	U26402 Rattus norv
43	153.4	3.9	1506	10	AF375461	AF375461 Rattus no
44	153.4	3.9	4654	10	RRP65	X52772 R. rattus p6
45	153.2	3.9	1461	10	AB026810	AB026810 Mus muscu

## ALIGNMENTS

RESULT 1  
ARI46667  
LOCUS ARI46667 3891 bp DNA  
DEFINITION Sequence 1 from patent US 6218523.  
ACCESSION ARI46667  
VERSION ARI46667.1 GI:15109856  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3891)  
AUTHORS French,C.K., Schneider,P.A. and Yamamoto,K.K.  
TITLE Prostate cancer-specific marker  
JOURNAL Patent: US 6218523-A 1 17-APR-2001;  
FEATURES  
source Location/Qualifiers  
1..3891

BASE COUNT 1229 a 671 c 788 g 1203 t  
ORIGIN

Query Match	100.0%;	Score 3891;	DB 6;	Length 3891;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3891;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Oy	1	ctcttggctctccctggcttcagcagagcctgfygsccttgagctcggcctgtgttttcaag	60
Db	1	ctcttggctctccctggcttcagcagagcctgfygsccttgagctcggcctgtgttttcaag	60
Oy	181	tttgatgaaatccccaagtggttgaggatcttcagtgcatlttgccctgtgtcttcacagtc	240
Db	181	tttgatgaaatccccaagtggttgaggatcttcagtgcatlttgccctgtgtcttcacagtc	240
Oy	241	tcctctcttgcatgatactgctgtgcagagaaatactccaagtcatacaagactctctcca	300
Db	241	tcctctcttgcatgatactgctgtgcagagaaatactccaagtcatacaagactctctcca	300
Oy	301	tacaagttctgtcagctgtcctaaagagttgtatcttaccctgtaaaaccctaaatagcaaa	360
Db	301	tacaagttctgtcagctgtcctaaagagttgtatcttaccctgtaaaaccctaaatagcaaa	360
Oy	361	aagaagttctggcgcagatgataaaatagaaataagaaataagccagactgtgcgcaagaat	420
Db	361	aaagaaattgttgagcgcagatgataaaatagaaataagaaataagccagactgtgcgcaagaat	420
Oy	421	tcattcatcttgagactctgtaaaagagagatctccaatgtgcaatttcccaaaaccaactc	480
Db	421	tcattcatcttgagactctgtaaaagagagatctccaatgtgcaatttcccaaaaccaactc	480
Oy	481	aaacctggagagcctcctcgtctcgtggagaaatgaaacccgaagccttttagaagggaa	540
Db	481	aaacctggagagcctcctcgtctcgtggagaaatgaaacccgaagccttttagaagggaa	540
Oy	541	aaagagtcagttccctcctgagagtttaaagtcacagcacttcccttactccaagaagaa	600
Db	541	aaagagtcagttccctcctgagagtttaaagtcacagcacttcccttactccaagaagaa	600
Oy	601	caagagaagcttggagact	660
Db	601	caagagaagccttggagact	660
Oy	661	gtgttcaatataaagaagcccgctgtgctctgtcagcaatctgtgacagtgatgtgacactt	720
Db	661	gtgttcaatataaagaagcccgctgtgctctgtcagcaatctgtgacagtgatgtgacactt	720
Oy	721	gaccctatatataaaatgaacatccctcccaagagaagaaactaaagtgaaactatagatg	780
Db	721	gaccctatatataaaatgaacatccctcccaagagaagaaactaaagtgaaactatagatg	780
Oy	781	cttgaagaaaacttgagatccagccttctgtgtgagaccttatactctctctctctctctct	840
Db	781	cttgaagaaaacttgagatccagccttctgtgtgagaccttatactctctctctctctctct	840
Oy	841	accctaaatccaaagattggcctctgtgcctcacaatttgaatttgaatttgaatttgaattt	900
Db	841	accctaaatccaaagattggcctctgtgcctcacaatttgaatttgaatttgaatttgaattt	900
Oy	901	gatgatatactctgtgggaagtctcaattctctctctctctctctctctctctctctctct	960
Db	901	gatgatatactctgtgggaagtctcaattctctctctctctctctctctctctctctctct	960
Oy	961	atgttaatagaatagagatcatacaagaagaaatgttaagaagtccttcaagacggggtgag	1020
Db	961	atgttaatagaatagagatcatacaagaagaaatgttaagaagtccttcaagacggggtgag	1020

[illegible]



Db	2101	AAAGTTACATGTGATGTCAAAATTTTGTAAATAATACCTCAATATAGAACCATGGCCTTG	2160
Qy	2161	gattattactgcgtgtcacaaagccctcag tgtgcgctgagaatccctatgtaccttgt	2220
Db	2161	GATTATTACATGGCCTGTACAAACGCTCACTGTGGCGCTGGAAATCCATGTACTTTGT	2220
Qy	2221	gaatgtgtaatatgtagttagtaataagaataaactcaactagaatccagtagaa	2280
Db	2221	GAATTTGTTAAATTTAGTTAGTATGATTAAGAAATAAATCACTCAACTGGAATCCAGTTAGAA	2280
Qy	2281	gtggaatttcctataggaatatagatgtggaaggtgaacttttaagccatgcgttt	2340
Db	2281	GTGCAATTTTCTTATATAGGAATATGATATGTGTGCAAGTATCTTTTAAGGCATCGTTT	2340
Qy	2341	gtacccagagtcgycgacatgagccacciaagcttcattcaatattatgtgccccagaanaa	2400
Db	2341	GTACCCAGAGTCGGCATGAGCCACACTTAAGTCTTCATTATTTATTTGTGCCCAAGAAAGA	2400
Qy	2401	ttaagatgctacttgtaaaagaacgtgtgaagatlltttaactgtgcagataaagaagtgtac	2460
Db	2401	TTAAGATGCTACTTGTAAAAACAGCTGTAAAGATTTTTCATTGCCAGATAAAGAGCTTAC	2460
Qy	2461	tttaaccaacaacaagaatgtaagaagctcaacaacgtttcaagaagcaattctaataat	2520
Db	2461	TTTAACCAACAACAATGTTATAGACTACAAATAAGTTCAAGACCAATTCATATATATATTA	2520
Qy	2521	catagtctcagcgaanaatatgctctagcgtgtccaatatgaacaacaagaatgtgtgttca	2580
Db	2521	CATATGTTCAACGCAAAATATGCTTTAGCGTCAAAATATACCAACAAGAAATGTGTTTCA	2580
Qy	2581	ctacacttcttagagctaaattgtctctgagcgtgtgtctctatagagcagtttcagaactgt	2640
Db	2581	CTACACTTTTCTAGCGTAAATTTGCTTGTGACCTGTGTCTATAGACGAGTTTACAGACTTGT	2640
Qy	2641	gtctctgtatcatllttccagcagcgaggtctctgaatatcttaagaacctgtttaagcttaa	2700
Db	2641	GTCCTGTATCATTTTCCAGTGGCCAGGGTTCTGAATTCATTTAGAACCTGTTAGATTAA	2700
Qy	2701	gctgcacccgtgtatctatttgaanaagaatctagctctgagaagtaatttcaactatattgt	2760
Db	2701	GCCTCACCCGTGTATTTTGGAAACAAATTTAGCTTGAAGTATCTCATATTTTGTAGT	2760
Qy	2761	tcttaagaagatlabagtvgaaacttgaagtcagttgaattatlaatatlgcaagltaga	2820
Db	2761	TCTTATAGGAAGTATGATGTGGAACCTTAGTATGAAATTTAATATATGCAAGTTAGAA	2820
Qy	2821	atlaagctcactgaaaaatttaactatttgaagtcaggttttgtgtcaagtaactttagcagt	2880
Db	2821	ATTAAGTCTACTGAAAAATTTTACATTTTGAAGTCAGGTTTGTGATGATCACTTTACCACTT	2880
Qy	2881	tttgaagaatgtgtgtatcaccagtggttgaataatcttgaanaaatgcaactttccaa	2940
Db	2881	TTTGAAGATGTGTTTGTATATACAGAGTGTTTGTATATCTATATAAATAATGCTATTTCCAA	2940
Qy	2941	caacttataatgcttlltatagtactatgccaatgtlaaagaanaatgltatlaactatctgt	3000
Db	2941	CAACTTATACATGCTTTTATATGACTATGCTATATGAAGAANAATGTATTACATCTGTA	3000
Qy	3001	tgtacaagaatlaaanaactcaactctlltttgtctttaaataatgactttggatlaaana	3060
Db	3001	TGTCAAGAAGATTAAAAATCAACCTCTTTTGTGCTTTAAAAATGACTTTGGATTAATAAA	3060
Qy	3061	agaaatattcccaatacttgttcaattccactaaagaatgcccacaagaatcttgtc	3120
Db	3061	AGCAATATTTCCATATCTATGTCTTCAATTCACATACAAATGCACTCACAGCATCTTGCTC	3120
Qy	3121	caactcggcatctcgttgaagaacaca tgaatgaactgaactgaagtagtgtatgttgggaa	3180
Db	3121	CACTCGGCATCTCTGTGAAGAACAACATGAATAATGAATCTGTATGAGTGTGTAGTTTGGGGA	3180
Qy	3181	agttcaaatggtccatttatagtatgtacgtatcgtatcgtggtcggtggaacagaatat	3240

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
3101	ACTCAAAATGGCCATT	TTTATGTATGTGCA	TTTGGATATCAT	TGGGCGCTGGACA	GCAGATATAT	3240				
QY	3241	gttgacctctgaaag	ttgtlaagggccaaatc	taaglatcttcacg	gcagccgaag	3300				
Db	3241	gTTGACCTCTGAA	AAGTTGTAAAGGGCCAA	ATCTAAATATTTCT	CACGGCAGCAAG	3300				
QY	3301	ttaatggtgttagc	gcgcgtgaagtagt	tglttgacgagccg	atltttttttaaca	3360				
Db	3301	TTTAAAGGNGTA	AGCAGCTAGAGTA	GTGTTGTGGCAGCGCCG	ATTTTTTTTTTAA	3360				
QY	3361	tggaaacaatgaa	ccaacaacaacacat	ctttttaaaatlaa	atgaatgtgaatagt	3420				
Db	3361	TGGAAATGGAAC	CAACAACAACACAT	TTTTTAAATTTGCA	TATTTGTAAATAGT	3420				
QY	3421	tttagctcttaaa	attaaagtgttttg	agtgtagaaagt	tgaglaaaactltg	3480				
Db	3421	TTTTAGCTTTTAA	ATTTTAAAGTGTG	TTTTGTGAGTGGA	AAAGTGTGAAC	3480				
QY	3481	actggttttcgaa	aaaggaagaaacaa	aaaggaaltga	aaacagcgagggat	3540				
Db	3481	ACTGTTTTTCGAA	AAAGGAAAGAAAGAA	ACAAACAAAGAA	ATTTGAACAGCAG	3540				
QY	3541	aatacctaattc	atcatcttcctgc	caaaatgacgt	ttttagaatgtat	3600				
Db	3541	AATACCTAATTT	CAATCATATTTCTG	CANAAATGTACT	TTTTTAGATGTAT	3600				
QY	3601	gtgaatcttly	aatccctgttaca	aatccctgcact	gtglaaacaatga	3660				
Db	3601	GTGAATCTTCTG	ATCTGTTTACA	AATCTGTGCTG	ATTTAAACATGT	3660				
QY	3661	tgctcgtatg	ccaatcttccacc	caaatatgggg	aggtatatact	3720				
Db	3661	TGCTGATTTAGC	CAATCTCACAC	CAACCAATGGG	AGGTATCATGTT	3720				
QY	3721	actcgtaatgat	ttctgtctcga	gtgttgaacta	atgaagtggtttg	3780				
Db	3721	ACTCGTAATGAT	TTCTGTTCTGAT	TTGTACTCAT	ATGAAGTGTTTG	3780				
QY	3781	ggtgtgtgag	acagtgctgtct	ttctgtgcacg	ctctgtatgalt	3840				
Db	3781	GGTGTGTGAGAC	AGTGCTGTTCTT	TGTGCGACGCTG	TATGATTTGT	3840				
QY	3841	ttgtgaagacat	gaaataatgctg	ctcttctgtcc	acgaataaaaaa	3891				
Db	3841	TTTGTAAAGCA	TCAGATTAAT	TGCTCTTTTG	CGCCAAAAA	3891				



|||||  
Db 1717 CTAGGAATGATGACGATTTTATGATTAATGCTGCCAGATCCCAATTAATATGTCACA 1776  
QY 1741 atctatgtgaacatgctcaaaaagttaatgtgattaagatttaaaacggaagatg 1800  
Db 1777 ATCTCAATGTGAACTGTCAAAAAAGTTAATGATTAAGATTTTAAAAAGAAATATG 1836  
QY 1801 ccttgctgtgaaaattatccattatcctcaagcttgagggaatacattttctta 1860  
Db 1837 CCTGGCTGTGAAAAATTTATCATTTATCTTCAGGTGGGCAATCAATTTTCTTTAA 1896  
QY 1861 atccaaagatacgaaaatgctccacagcttgatattatattatcgtgcatgtaaa 1920  
Db 1897 ATCCAAAGATTAATAAAAAATGCTCCAGTTGTGATTTATTAATCTGTCATGTGCAAA 1956  
QY 1921 tgggtgctgcatataaagaatcgtgcattcctcagcttggttgatttaatttgatgc 1980  
Db 1957 TGGTGTCTCGATATTAATAAGATGTCATTTGCTGATTTGTTGTTGATTTATTCATGC 2016  
QY 1981 aatttatcataagatacgaattcgaattcctcaaaagacagtgacaagctggaat 2040  
Db 2017 AATTTTATATAGAGTAACTGATTCATTTCAAAAGACAGTGAACAAGCTGAGAAAT 2076  
QY 2041 taattatcaaaaggctgaagctgaagacacgtgagctgaataataattttcccccct 2100  
Db 2077 TATTTTATCAAAAGGCTGAGTTGAGAACACTGCGCTGAATATATTTTCTCCCCCT 2136  
QY 2101 aaggttacaatgtgagtcgaatattgttaataataaccaccaagaacatgacctg 2160  
Db 2137 AAGGTACATGTGAGTCMAAAATTTGTAAATAATACCTCACATAAGAACCAAGGCTTG 2196  
QY 2161 gattatcactgcgcgtgcaagcctcagctgagcctggaataccctatgtaacctgt 2220  
Db 2197 GATTAATTCACCTGCTGTCAAAAGCCTCATGTGCTGGAATTCCTTATGTACCTTTGT 2256  
QY 2221 gaaatgtgtgaattagctgaatgaataaagaataaactcaactagaanaaccagctaga 2280  
Db 2257 GAATGTGTGAATGATGATGATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 2316  
QY 2281 gtcgaatttctctatagaanaatagatagctgtgcaagctgactcttaaggccaactgt 2340  
Db 2317 GTGCAATTTTCTTATAGGAATAGATGATGATGATGATGATGATGATGATGATGATG 2376  
QY 2341 gtaccagaagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400  
Db 2377 GTACCCAGAGTGGGCTGAGCCACCTAGCTCATTTAATTTATTTCCGCCGAAAGAA 2436  
QY 2401 ttaagaatgactggaagaagctggaagaatttttacaattgcagataaaaaagctgac 2460  
Db 2437 TTAAGATGCTACTGAAAAGACGTGAAGATTTTACATTCGCATATAAAGTTTAC 2496  
QY 2461 ttaaccaacaacaaatgtaagaactaanaacgltcaagaagcaattctaataatla 2520  
Db 2497 TTAACCAACAAACAAATGTAGACTCAAAATCGTTCAAGAGCAATTTCAATTAATTTA 2556  
QY 2521 catatgttacaagaaatgctgaagctgcaaatlaagcaaaaagaatgtgttca 2580  
Db 2557 CATATGTTACGAAATATACTTATGAGCTGTCAAAATAGACACAAAGATGTGTTCA 2616  
QY 2581 ctatctttctgaagctgaattgctgaagctgtgtcctatagaagcagcttcaagactgt 2640  
Db 2617 CATATCTTTCTAGGCTAATTTGTGCTGAGCTGTGTCTATAGAGCAAGTTTACAGACTGTG 2676  
QY 2641 gctctgtatcaatttccagctgccagagcttcgaaatcatcagaacctgttagatlaa 2700  
Db 2677 GCTCTGTATCATTTTCCATGCGCAGGCTCTGAAATTCATTCAGAACCTGTGTAGTTAAA 2736  
QY 2701 gctgacccctggaattatttgaagaagaattagcttgaagataatgcacataattgaat 2760  
Db 2737 GCTGACCCCTGTGATTTATTTGAAGAATTTAGCTTGAAGATTAATGACATATTTGAGT 2796  
QY 2761 tcttgaagaagatagctgaagcagctgaagctgaagctgaagctgaagctgaagctgaag 2820  
|||||

Db 2797 TCTTAGAAGATGATGAGGAACTGTAGTACAGTTGATTAATTAATGACAGTTAGAA 2856  
QY 2821 attaaatctacgaaaaatttaccatttgaagctgttgcagtaacttaagcaat 2880  
Db 2857 ATTAAGCTTACGAAAAATTTACATTTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2916  
QY 2881 ttgaagaatgtgttgaatacgaagctgttgaatacttcaatgaagaatgactttccaa 2940  
Db 2917 TTTGAGAATGTGTTTATATACAGAGTGTGTAATTTATGAAATAATGATTTTCCAAA 2976  
QY 2941 caactatacatgcttcttatagtactagccaaatgaagaanaatgataltaactgtat 3000  
Db 2977 CAACTTATCATGTTTATGACTATGCTAATGTAAGAAATGATTAATTCATTCGTA 3036  
QY 3001 tgfacaagaatlaaaaatacactcttttggctttaaagactttggagtttaaaa 3060  
Db 3037 TGACAAAGATTAATAATCACTCTTTTGTGCTTTTAAATGACTTTGGGATTAATAA 3096  
QY 3061 agcatacttcccaatcatgtcttctatccactacaagaatgcacccaagcactgtgctc 3120  
Db 3097 ASCATTTTCCCAATCATTTGCTTCATTCACATCAAGACACTCAGACATCTGCTC 3156  
QY 3121 caactggcatctctgtgaagaacacatgaatgaactgtagtgctgtagcttgaggaa 3180  
Db 3157 CACTCGCATCTCTGTGAAGAACAATGAATGAATGAATGAATGAATGAATGAATGAATGA 3216  
QY 3181 agtcaaatggcaattatgatagtgcatgttgatcatgagcgcgtggagaagaatata 3240  
Db 3217 AGTCAATGGCCATTTATGATGTGATGATGATGATGATGATGATGATGATGATGATG 3276  
QY 3241 gctgagccctggaagaatgtgaaggcccaatcgaatcttcaagcgaacgaag 3300  
Db 3277 GTTGGACCTCTGAAAGTTGTATAGGGGCCAATTCATGATTTCTTCCAGCGACCCGAG 3336  
QY 3301 ttaatgtgtgaagcagctgaagatagctgtgtgagcagagccgaattttttttaa 3360  
Db 3337 TTAATGCTGTAGTACACTGAGTATGTTGTGAGAGAGCCGATTTTATTTTAAACA 3396  
QY 3361 tggaaacaatgaaacaaacaaacaaatlttaaatlaaatgataatttgaataatg 3420  
Db 3397 TGGAAACATGAACCAACAAACAAATTTTAAATTAATTAATTAATTAATTAATTAAT 3456  
QY 3421 tttaagctttaaatttlaagtgcttgaagtgtaaaagtgaagtaaacatttgc 3480  
Db 3457 TTTTACTTTTAAATTTTAAAGTGTTTTGAAGTGTGTAAGAAAGTTGAGTAAACATTTGCA 3516  
QY 3481 actggttttcaagaagaagaagaagaagaagaatgaaacagagcagagagactc 3540  
Db 3517 ACTGTTTTCAGAAAAGAGAAAAGAAAACAAAGAAATTTGAACAGGCGAGATCTT 3576  
QY 3541 aatacctaatttcaatctctgcaaaatgactgttttagaatgtatcaatcaat 3600  
Db 3577 AATACCTAATTTTATCATTTTCTG-AAAATGTACTGTTTGAATGTATTAACAATATCAAT 3635  
QY 3601 gtaaatatcttgaatcctgttacaacacccgcaatlatlaaacaatgaataattgt 3660  
Db 3636 GTCAATATCTTGAATCTGTGTACAAATCTCGACCTGATTAACATGAATTAATTTGTT 3695  
QY 3661 tgcctgattagccaatctcaaccacaaatgaggaggtatatacttgaagaac-gtgt 3719  
Db 3696 TGTCTGATTAAGCAATCTCACCCCAAAATGGGAGGTATACATGTTGAAGAACTGTG 3755  
QY 3720 aactcgttaattgaattgtctctgaatgttgaactcaatagaatgttttgaagaagca 3779  
Db 3756 AACTCAGTAAATTAATTTGTTGATGTGTGAATCAATTAAGATGTTTGAAGAAAGCA 3815  
QY 3780 tgggtgtgaagacagctgctcttcttggcagcctctgtagaagtttgaagacat 3839  
Db 3816 TGTGTGTGAGACAGTGTCTGTTTGTGTCACGCTGTATGATGATGTTGTGAAGACAT 3875  
QY 3840 gtttgaagacatgaaataattgtcgttcttggcc 3874  
Db 3876 GTTGTGAAGACATGAATTAATTTGCTGCTTTTGGCC 3910  
|||||

RESULT 3  
AC009266/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AC009266.117177 bp DNA HNG 09-MAY-2001  
Homo sapiens chromosome 18 clone RP11-481E14 map 18, WORKING DRAFT  
SEQUENCE, 4 unordered pieces.  
AC009266  
AC009266.4 GI:13112247  
HNG: HNGS\_PHASE1; HNGS\_DRAFT; HNGS\_FULLTOP.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 17117)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 18, clone RP11-481E14  
Unpublished  
2 (bases 1 to 17117)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baker,J., Baldwin,J., Barna,N., Beckert,J.R., Benn,J., Brown,A.,  
Castle,J., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,P., Derrallano,K., Depayre,E., Devon,K., Dewar,K.,  
Donald,J., Doyle,M., Ferreira,P., FitzHugh,M., Forrest,C.,  
Fung,E.R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,  
Hague,B., Haefford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,  
Kardas,A., Lebeckzy,J., Lieu,C., Locke,K., Macdonald,P.,  
Marais,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,  
Melartin,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,  
Naylor,J., Nijhoff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,  
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,  
Thesfaye,S., Tortorella-Miller,I., Vassiljev,H., Vo,A., Wagner,A.,  
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.  
Direct Submission  
Submitted (11-APR-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 23, 2001 this sequence version replaced gi:724941.  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green,P (1996)1997  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WtBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
Project Information  
Center project name:11029  
Center clone name:481.E-14  
Summary Statistics  
Sequencing vector: M13, 477815; 62% of reads  
Sequencing vector: Plasmid, n/a; 38% of reads  
Chemistry: Dye-Primer-amer.shim; 4% of reads  
Chemistry: Dye-terminator Big Dye; 96% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 169442 bases at least Q40  
Consensus quality: 170635 bases at least Q30  
Consensus quality: 171095 bases at least Q20  
Insert size: 176000; agarose-fp  
Insert size: 171417; sum-of-contigs  
Quality coverage: 6.5 in Q20 bases.  
NOTE: This is a 'working draft' sequence. It currently  
consists of 4 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 39315: contig of 39315 bp in length  
\* 39316 39415: gap of 100 bp  
\* 40721 40720: contig of 1305 bp in length  
\* 40721 40820: gap of 100 bp

[illegible]

Db 118143 AGAATCCAAATTAATATGTGACAATCTCANTGNAACATGTCACAAAAGTTAAATGTGA 118084  
QY 1276 ttaagatttaaaagaaagatgcttgcttgcttggaatatcttatctctcag 1835  
Db 118083 TTAAGATTTTAAAACGAAAGATGCTTGCCCTTGGAATAATTTATCAATTAATCTTCAG 118024  
QY 1836 gtctgggagaaatcaatcttctcttaacaaagatacctaataaaatgtccctcagtttg 1895  
Db 118023 GTTGGGAATCAATTTTCTTTAAATCCAAAGATATCAAAAAATGTCCTCAGTTTGT 117964  
QY 1896 attataatctgcatactgcaaatgltgcctgcataataaaglatcgtgcatctc 1955  
Db 117963 ATTATTAATCTGTCATGTGCAAAATGTTGTCCTCATATTAAGATATCTGTCATTTTC 117904  
QY 1956 agtttggttgtaattatttgatgtcaatttatcataaagtaacctcagatcttcaaa 2015  
Db 117903 AATTGGTTTGTAAATTAATTTGATGCAATTTTATCAATTAAGATATCAATTCATTTTCA 117844  
QY 2016 aaggacagtgaaacaaagctgagaatatttlatcaaaaggctgagtggaacactgttg 2075  
Db 117843 AAGGACAGTGAACAAACGTCAGAAATTTATTCAAAGGCGCTGAGTGAACACTGTG 117784  
QY 2076 ctgaaataataatttctcccccctaaaggttacaatgtagcaaaatlctgttaaaatata 2135  
Db 117783 CTGAATATTAATTTTCTCCCTTAAGSTTACATGTGAGTCAAAATTTTGTAAATATATA 117724  
QY 2136 acctacataaagaacaaagcttgatattatctcactgctgtgcacaaagctcagtgctgc 2195  
Db 117723 ACCTACATAAAGAACGCTGATGATTTTACATGCTGACAAAGCTCAGTGTG 117664  
QY 2196 ctgagaatccctatgtaaccttgtaaaatlctgtaactagtaggaataaagaataa 2255  
Db 117663 CTGAGAAATCCCTATGTACTTGTGAATTTGTGATTAAGTATGAATTAAGAAATATA 117604  
QY 2256 acctcaactagaatccagtgtaagtgcaattctctatagaanaataggtatagtgctc 2315  
Db 117603 ACTTCAACTAGAAATCCCAATTTGAAGTGCATTTTCTTATAGGAATTAAGTATGAGTGC 117544  
QY 2316 aagtgtaactttaagagcagtggttgaccagagctgagcagcactaagctctcact 2375  
Db 117543 AAGTGACTTTTAAGGCCATGCTTGTACCCAGAGTCGGCATGCGCACTAAGCTTTCAT 117484  
QY 2376 ttaatttatgtcccccgaataaagaltgaagatgtaactgtaaaagactgtaagattctc 2435  
Db 117483 TTAATTTATGTGCCCCCAAAAAGATTAGATGCTTGTGAAAAGCTGTGAAGATTTT 117444  
QY 2436 taacttgccaagataaagtgtaacttaaccaaacaacaatgtaagactacaaatcgtc 2495  
Db 117423 TACATTGCCAGATTAATAAAGTGTAACTTAACCAACAATAATGTAAAGACTACAAAATCGT 117384  
QY 2496 taaagagcaacttaataattacatagttcaagcgaataatgcttagctgtcaca 2555  
Db 117363 TCAAGAGCAATTTCTAATTAATTTACATATGTTCAAGCAAAATATGCTTAAGCTGTCA 117304  
QY 2556 ttagacaacaagaatggtgttcaactatcttcttagagctaaatggtcttgagctgtg 2615  
Db 117303 TTAGCAACAACAAGATGTGTTCACTATCTTTTCAAGCTAATTTGTCTGAGCTGTG 117244  
QY 2616 tctataagagcagtttaacagactgtgtctgtcatcaatcttccagtgccagggctcga 2675  
Db 117243 TCTAATAGACAGTTTAACAACACTGTGTCTTGTATCATTTTCCAGTCCAGGCTTCCAAA 117184  
QY 2676 ttcattcaaacctttagaattaaagctgcaccctgtaattttgaagaatagctct 2735  
Db 117183 TTCATTCAAACTGTTAATTAATTAAGCTCACCTGTGATTTATTTGAAAAGATTAAGCTT 117124  
QY 2736 gagaagtaagtcatactatctgagttcttagagaagatagagtggaagactgtgacagtt 2795  
Db 117123 GAGAGTAAGTCACTATATTAGTCTTAGAGAACTATGAGTGAACCTTGATGATACGTT 117064  
QY 2796 gaattataataatgcaagtgagaataagtcactactgaaaaatctacatcttgagtcag 2855  
Db 117063 GAATTATTAATAATATGCAAGTTAGAAATTAAGTCTACTGAAAATTAATTTTGAAGTCA 117004

QY 2856 gtcttggtcagtaacttagcagtttttgagaatggtgttgatatacacagtggtttaa 2915  
Db 117003 GTTTTGTGCACTACTTTAGCAGTTTGTAGAGATGTTGATATACAGTGTGTTTAA 116944  
QY 2916 tctatgaanaaalgcatcttccaaacaactatacagctcttlatagactatgcctaalg 2975  
Db 116943 TTCTATGAATAATGCAATTTTCCAAACACTTATACATGCTTTTATGACATGCTCAAG 116884  
QY 2976 taaagaanaatgtaatacattcgtatgtaacaagaatlaaanaatcaactcttcttgctgc 3035  
Db 116883 TAAAGAAATGATTTATCATCTGTATGTACAAAGATTAATAAATCAACCTTTTGTGTC 116824  
QY 3036 ttttaaatgactttggagttaaanaagcalttcccaatcactgtctcattccactac 3095  
Db 116823 TTTAATAATGCTTTGGATTAATAAAGCAATATTTCCCAATCATTTGCTTCATTTCCACTAC 116764  
QY 3096 aaagtcacctcacagatctgtccactcagcatctcgtgaaagcaacatgaatga 3155  
Db 116763 AAAGTCACTCACAGCATCTGCTCCACGCGCATCTGTGAAACCAATGAAATGA 116704  
QY 3156 ctgtagtaggtgtgtagtttggygaagtcacaatgycatcttaatgtagtgcaattgta 3215  
Db 116703 CTGTAGTAGGTGTGTAGTTGGGGAAGTCAAAATGCCATTTTATGTATGTGCAATTTGTA 116644  
QY 3216 taatggccgttgaaacagaatattgtgacctcgtgaanaagtgtaagggccaaatct 3275  
Db 116643 TCATGGCCGTGGAACGATATATATGTGACCTCGAAAAGTTGTAAGGGCCAAATTC 116584  
QY 3276 aagtaattctcaagcagcagaagtaatgtygtgaagcagctgagtgtagtggtga 3335  
Db 116583 AAGTATTTCTACAGCAGCAGCAGAGATTAATGTGTGATGACGCTGAGATATGTTGGA 116524  
QY 3336 cgaagccgattttttttttaaactgaaacaaatgaacaaacaaacattttaa 3395  
Db 116523 CGAGGCCGATTTTGTTTTAAATGAAACATGAAACCAACCAACATTTTAA 116464  
QY 3396 ttaaaatgataatgtgtaaatgattttagctctttaaataatgaatgttcttgagt 3455  
Db 116463 TTAATAATGATTAATTTGTAAATATGTTTAAATGCTTTTAAATTTTAAAGTGTGAGTGT 116404  
QY 3456 gaaaagttagtaaaactatctgcaactgtgttccagaagaagaagaacaaacaaag 3515  
Db 116403 GAAAAGTTGATGAACCTATTTGCAACGCTGTTTGAAGAAAAGAAACAAACAAAG 116344  
QY 3516 gaattgaacagcagcagcagccttaatacctaattcatcatcttgcgaatgtactg 3575  
Db 116343 GAATTGAACAGCAGCAGGAGATCTTAATACCTTAATTCATCATTTCTG-AAAATGTACTG 116285  
QY 3576 ttttagaatgtaataataatgataatgtaataatctgtaactcgtgtacaaatcctcact 3635  
Db 116284 TTTTAAATGTATTAACAATATCAATGTGAATATCTTGAATCCGTTRACAAATCTCACT 116225  
QY 3636 gtaataacatgtaaatgtaattgtgtcgtatgtaacatctcaccaccacaatggga 3695  
Db 116224 GTATTTAAACATGAATTAATTTGTTGTCGATTAACCAATCTCACACCAAAATGGGA 116165  
QY 3696 ggtatacatglttgaagaac-gtgtactcogtgaattgattgttcttgatgtgtaacc 3754  
Db 116164 GGTATTCATGTTTGAAGAAGCTGTGAACACAGTAATTTGATTTGTGATGTTGTATC 116105  
QY 3755 aatagaagtggttgaagaagcagtggtgtgtaagacagtgctgtcttcttggtgcag 3814  
Db 116104 AATGAAGTGTTTGGAAGAGCAATGCTGTGAGACAGTGTCTTTCTTTGTGTCAG 116045  
QY 3815 ctctgtatgattgttgtaagaccatgttgtaagacatgaataaattgctgtcttgccc 3874  
Db 116044 CTCGTATGATGTTTGTAAAGACCATGTTTGTAAAGACATGAATAATGCGCTTTGCC 115985  
QY 3875 aa 3876  
Db 115984 AA 115983

RESULT 4  
AP001797 190550 bp DNA HTG 30-MAY-2000  
LOCUS Homo sapiens chromosome 18 clone RP11-748120 map 18q12, WORKING  
DEFINITION DRAFT SEQUENCE, 27 unordered pieces.  
ACCESSION AP001797  
VERSION AP001797.2 GI:8117469  
KEYWORDS HTG; HTGS; PHASE1; HTGS-DRAFT.  
SOURCE Homo sapiens DNA, clone RP11-748120.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 190550)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.  
TITLE Homo sapiens 190,550 genomic DNA of 18q12  
JOURNAL Published Only in Database (2000) In press  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (11-Apr-2000) Masahiro Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
Katsushika Univ., 1-15-1 Katsushika, Sagamihara, Kanagawa 228-8555,  
Japan (E-mail:hattori@gsc.riken.go.jp,  
URL:http://hbp.gsc.riken.go.jp/, Tel:81-42-778-9923,  
Fax:81-42-778-9924).  
On May 30, 2000 this sequence version replaced gi:7592912.  
COMMENT ----- Genome Center  
Center: RIKEN Genomic Sciences Center(GSC)  
Center code: RIKEN  
Web site: http://hbp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: HumPrat18  
Center clone name: RP11-748120  
----- Summary Statistics  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye terminator; BigDye; version 0.390329  
Assembly program: Phrap; version 0.390329  
Consensus quality: 188743 bases at least Q40  
Consensus quality: 188743 bases at least Q30  
Consensus quality: 188891 bases at least Q20  
Insert size: 187950; sum-of-ctrls  
Quality coverage: 4.84x in Q20 bases; sum-of-ctrls  
NOTE: This is a 'working draft' sequence. It currently consists of  
27 contigs. The true order of the pieces is not known and their  
order in this sequence record is arbitrary. Gaps between the  
contigs are represented as runs 'N', but the exact sizes of the gaps  
are unknown. This record will be updated with the finished sequence  
as soon as it is available and the accession number will be  
preserved.  
1  
16819 16718 contig of 16718 bp in length  
30449 contig of 13631 bp in length  
30550 45607 contig of 15058 bp in length  
45708 55752 contig of 14045 bp in length  
55853 75894 contig of 16042 bp in length  
75995 87034 contig of 11040 bp in length  
87135 97062 contig of 9928 bp in length  
97163 106435 contig of 9273 bp in length  
106536 114800 contig of 8265 bp in length  
114901 121600 contig of 6700 bp in length  
121701 129422 contig of 7722 bp in length  
129523 138872 contig of 9350 bp in length  
138973 145365 contig of 6393 bp in length  
145466 151207 contig of 5742 bp in length  
151308 155673 contig of 4566 bp in length  
155774 160475 contig of 4702 bp in length  
160576 165680 contig of 5105 bp in length  
165781 169730 contig of 3850 bp in length  
169731 174769 contig of 5039 bp in length  
174770 174869 contig of 100 bp  
174870 178072 contig of 3203 bp in length  
178073 178172 contig of 100 bp  
178173 180351 contig of 2179 bp in length  
180352 180451 contig of 100 bp  
180452 181295 contig of 844 bp in length  
181296 181395 contig of 100 bp  
181396 184552 contig of 3157 bp in length  
184553 184652 contig of 100 bp  
184653 186282 contig of 1630 bp in length  
186283 186382 contig of 100 bp  
186383 188196 contig of 1814 bp in length  
188197 188296 contig of 100 bp  
188297 189386 contig of 1090 bp in length  
189387 189486 contig of 100 bp  
189487 190550 contig of 1064 bp in length.  
Location/Qualifiers  
1. 190550  
/organism="Homo sapiens"

174870 178072 contig of 3203 bp in length  
178173 180351 contig of 2179 bp in length  
180452 181295 contig of 844 bp in length  
181396 184552 contig of 3157 bp in length  
184653 186282 contig of 1630 bp in length  
186383 188196 contig of 1814 bp in length  
188297 189386 contig of 1090 bp in length  
189487 190550 contig of 1064 bp in length  
Sequence updated (26-May-2000).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 27 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of 'N', but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.  
1  
16718 contig of 16718 bp in length  
16819 16818 gap of 100 bp  
16819 30449 contig of 13631 bp in length  
30450 30549 gap of 100 bp  
30550 45607 contig of 15058 bp in length  
45608 45707 gap of 100 bp  
45708 55752 contig of 14045 bp in length  
55753 55852 gap of 100 bp  
55853 75894 contig of 16042 bp in length  
75895 75994 gap of 100 bp  
75995 87034 contig of 11040 bp in length  
87035 87134 gap of 100 bp  
87135 97062 contig of 9928 bp in length  
97063 97163 gap of 100 bp  
97163 106435 contig of 9273 bp in length  
106436 106435 gap of 100 bp  
106436 114800 contig of 8265 bp in length  
114801 114800 gap of 100 bp  
114901 121600 contig of 6700 bp in length  
121601 121700 gap of 100 bp  
121701 129422 contig of 7722 bp in length  
129423 129522 gap of 100 bp  
129523 138872 contig of 9350 bp in length  
138873 138971 contig of 930 bp in length  
138973 145365 contig of 6393 bp in length  
145366 145465 gap of 100 bp  
145466 151207 contig of 5742 bp in length  
151208 151307 gap of 100 bp  
151308 155673 contig of 4566 bp in length  
155674 155773 gap of 100 bp  
155774 160475 contig of 4702 bp in length  
160476 160575 gap of 100 bp  
160576 165680 contig of 5105 bp in length  
165681 165780 gap of 100 bp  
165781 169730 contig of 3850 bp in length  
169731 169730 gap of 100 bp  
169731 174769 contig of 5039 bp in length  
174770 174869 gap of 100 bp  
174870 178072 contig of 3203 bp in length  
178073 178172 gap of 100 bp  
178173 180351 contig of 2179 bp in length  
180352 180451 gap of 100 bp  
180452 181295 contig of 844 bp in length  
181296 181395 gap of 100 bp  
181396 184552 contig of 3157 bp in length  
184553 184652 gap of 100 bp  
184653 186282 contig of 1630 bp in length  
186283 186382 gap of 100 bp  
186383 188196 contig of 1814 bp in length  
188197 188296 gap of 100 bp  
188297 189386 contig of 1090 bp in length  
189387 189486 gap of 100 bp  
189487 190550 contig of 1064 bp in length.  
Location/Qualifiers  
1. 190550  
/organism="Homo sapiens"

FEATURES  
source

```
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12"
/clone="RP11-748120"
1..16718
misc_feature
/note="assembly_fragment"
16819..30449
/note="assembly_fragment"
30550..45607
misc_feature
/note="assembly_fragment"
45708..59752
misc_feature
/note="assembly_fragment"
59853..75894
misc_feature
/note="assembly_fragment"
75955..87034
misc_feature
/note="assembly_fragment"
87135..97062
misc_feature
/note="assembly_fragment"
97163..106435
misc_feature
/note="assembly_fragment"
106536..114800
misc_feature
/note="assembly_fragment"
114901..121600
misc_feature
/note="assembly_fragment_clone_end:SP6_vector_side:right"
121701..129422
misc_feature
/note="assembly_fragment"
129523..138872
misc_feature
/note="assembly_fragment"
138973..145365
misc_feature
/note="assembly_fragment"
145466..151207
misc_feature
/note="assembly_fragment"
151308..155673
misc_feature
/note="assembly_fragment"
153774..160475
misc_feature
/note="assembly_fragment"
160576..165680
misc_feature
/note="assembly_fragment"
165781..169630
misc_feature
/note="assembly_fragment"
169731..174769
misc_feature
/note="assembly_fragment"
174870..178072
misc_feature
/note="assembly_fragment"
178173..180351
misc_feature
/note="assembly_fragment"
180452..181295
misc_feature
/note="assembly_fragment_clone_end:T7_vector_side:right"
181396..184552
misc_feature
/note="assembly_fragment"
184653..186282
misc_feature
/note="assembly_fragment"
186383..188196
misc_feature
/note="assembly_fragment"
188297..189386
misc_feature
/note="assembly_fragment"
189487..190550

Query Match 69.9%; Score 2720.4; DB 2; Length 190550;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2749; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
```

```
QY 1296 gtcccgaaatgagtaatcggcagtgatcttgggtgcagcagcgaagaagcagtggtg 1355
DB 13278 GTCCCGAAATGAGTAATCGGGCAGTTAGCTTGGGTGAGCAGCAGAAAGAACTGGTGG 13337
QY 1356 agagcactgaaagagatctgtaactcccgaggaacaaattgccaagtgcagtgat 1415
DB 13338 AGAGCAGCTGAAAGAGATCTGTGACTACCCCGAGAGACAAATTGCCAAGTGCACGTGCT 13397
QY 1416 ctgtgagtgtagcatccctagccgtgagttggaacttaaggatttactaagcaaggag 1475
DB 13398 CTGTGATGTTAGCATCCAGCCGTGAGTTGGAACCTTAAGGTTTACTAGCAAGAGAG 13457
QY 1476 aaatttcttcttctatattggaatgcagctggggaaatccaagctacatttgggtg 1535
DB 13458 AAATTTCTTCTTCTTATATTGGATTGCAACCTGGGAAATCAAGCTTCCTTTGGTTGG 13517
QY 1536 ttgtgtgttgctagaatgagttgaattagtagaccgaagaagtaactcaatgtgta 1595
DB 13518 TTGTTGTTTGTGCTAGTAATGATTTGAATTAGTACCGAAAGTAACCTCAAAATGTGTA 13577
QY 1596 ctatgtaattccctatttattagaagagttggabaatttccataagatatcaatat 1655
DB 13578 TTATGATTAATTTCCCTATTATTAGAGAGATTGGAATAAATTTTCATTAAGATTCGAATAT 13637
QY 1656 ctccctcaattaccagtgatataactaaggaatagtcagacatttataactgtgccc 1715
DB 13638 CTCCCTCAATTAACAGTATTAATAGTAGAATGACACATTTTATTAATATCTGTGCC 13697
QY 1716 agaattcccaattataaagtgtaaccaatcccatgtgacactgtccaaagaatgaatgtga 1775
DB 13698 AGAATCCCAATTATTAATAGTGACATCTCATTTGSAACATGTCACAAAAGTTAATGTGA 13757
QY 1776 ttaagattcaaaagaaagaaagatagtccttgccctgtgaaattatccaattacttcag 1835
DB 13758 TTAAGATTAAAAACCAAAAGATGCTGCTTGGAATAATTTATCCATTAATCTTCAG 13817
QY 1836 gtgggggaaatcaaatcttctttaaaccagaagatctaaaaaagtctccagttgt 1895
DB 13818 GTTGGGAAATCAATTTTCTTTAAATCCAAAGATATCTAAAAAAGTCTCCAGTTTGT 13877
QY 1896 attataatctcgtcatgtgcaaatgtgtgctgcataataaagtatcgtgcatctc 1955
DB 13878 ATTTATTATTTCTGTCATGTGCAAAATGTTGCTCCATATAAAGTATCTGGTCAATTC 13937
QY 1956 agtttggttctgaattatcttgatgcaatttatcataaagaatcaatccatcttcaa 2015
DB 13938 AGTTTGTTGTGAATTAATTAATGCAATTTTATCAATGAAGTAACACAGATTCATTTCAA 13997
QY 2016 aaggacagtgaacaagctgagaagaatttatcaaaagggctgagttggaacactgtg 2075
DB 13998 AAGGACAGTGAAACAGCTAGAAATTAATTTATCAAAAGGGCTGAGTGGAACACGTGCG 14057
QY 2076 ctgaaatataattttctcccccctaaggttacatgtgagtcacaaaatttgytaaaatata 2135
DB 14058 CTGAATTAATAATTTTCTCCCCCTTAAGGTTACATGTGAGTCAAAATTTGTGAATATA 14117
QY 2136 acctcaataaagaacacatgagcttgatatactacgtcctgtgcaaaagcctcagttgag 2195
DB 14118 ACCTCAATGAAGAACCATGAGCCTTGATTAATACACGCTGTCAACAAAGCTCAGTGTGAC 14177
QY 2196 ctgagaagaatcccatgtaccttggtgaatgtggaattagttagttggaataaagaataa 2255
DB 14178 CTGAGAAATCCCTATGTACTTTGTGAATAATTTGTTGAATTAAGTAAAGAAATATA 14237
QY 2256 acctcaactagaatcccaagttagaagtgaatttcttataaggaataagttatagtgctc 2315
DB 14238 ACTTCAACTAGAATCCAGTTGAAGTGAATTTCTTAAGGAATTAAGTATAGTGTGC 14297
QY 2316 aagtgtaactttaaggcatcggtgtgtaaccagagtcgggcaatggccaactaagttcatt 2375
DB 14298 AAGTGTACTTTTAAGGCATCGTTGTGTACCCAGAGTCGGCATGGCCACTAAGTCTTCAAT 14357
```



OY	2376	ttaattatatttcccccaaaaagattlaagaatgctacttgaanaagctgtgaagaattttc	2435
Db	14558	TTAAATTATGTCCCCAGAGAAAGATTAAAGATGCTACTGTAAGAAAGCTGTGAAGATTTT	14411
OY	2436	tacattgcacagataaaaagtgttacttcaaccaacaacaagttaagactacaatcgtc	2495
Db	14418	TGATATTGCCAGTATTAAGATGTTACTTTAAACCAACAACAATGATTAAGACTATCAAAATCGT	14477
OY	2496	tcaagaagcaattctaatatataatttaacatagtctcaagaaataatagcttagctgtc	2555
Db	14478	TCAAGAGCAAAATTCTAAATATATATTACATATGTTCACCCAAATAATCTTTAGGCTGTCCAA	14531
OY	2556	ttagccaacaacaagaatagtcttccactactctctcaggtctaaattgtcttgcagctgtg	2615
Db	14538	TTAGGCCAAACAAGAAAGTGTTCACATATCTTTCTAGGCTAAATTTGTCTGTAGCTGTGG	14597
OY	2616	tccatagagcagtttcaagaagctgtgtctctgataatttccaggtccagggctgtgaa	2675
Db	14598	TCTATTGAGCAAGTTTACAGACTGTGTCTGTATCATTTTCCAGTCCCGAGGGTCTGTAA	14655
OY	2676	ttaattccagaacccgttagaattaaagctgcaccctgcagatatttgaanaagatattgct	2735
Db	14658	TTTCATTGACAACCTGTATGATTAATTAAGCTGCACCCGTGATTAATTGTAAGAAATTAAGCTT	14711
OY	2736	gaagaagaatgcactatattttagtttcttaagaagaatataagagactaagtaagct	2795
Db	14718	GAGAGTAATGCTACTATATTGTGAGTCTTTAGAGAAAGTATGAGTGAAGACTTGAGTACAGTT	14777
OY	2796	gaatttaataataagcaagttaaaaaattaaatctacagaaaaatttaacttttgaagcag	2855
Db	14778	GAAATTTTAAATATGCAAGTTAAATTAAGTCTACGTAAATAATTTACATTTTGTGCTAG	14833
OY	2856	gtttctgtcagtaactttagcagcttttggagaagtgtctgataccaagtgctgtgaa	2915
Db	14838	GTTTTGTGCAATCTCTTTTACAGAGTTTGTAGAAAGTGTGTGATATACAGGTTTGTGAAA	14897
OY	2916	tctcatgaaaaagcgaatttccaaacaactatacatagcttllatgaactlgtcctaag	2975
Db	14898	TTCTATAGAAAATGCAATTTTCCAAACAATTAATATACATGCTTTTATGACTATAGCCTTAATG	14957
OY	2976	taaagaanaatgatcatcattctgtatgtcaagaagataaaaaatcaaacctcttttctg	3035
Db	14958	TAAAGAAATGATTAATTAATTTCTGTATGTACAAAGATTAAATAATCAACCTTTTGTGTGC	15011
OY	3036	tttaaaatgaacttgggattaaaaaagaactatcccaacatctgtctctatccaactc	3095
Db	15018	TTTTAAATATGACTTTGGGATTAAAAAAGCATATTTCCCAATCATTTGTCTTATCCACTAC	15077
OY	3096	aaagtlacccaagaacatctgtcccaactcggacatctcgttgaanaacaacatgaaatgaa	3155
Db	15078	AAAGTACTACTCAGCAAGCATCTTTCACCTCAGCGCATCTCTGTGAAGCAACATGAAATGAA	15133
OY	3156	cgtgtagtagtgcttaatttggggaagtcaaatgagccattttagatgtgatttgta	3215
Db	15138	CTGTAGTAGTGCTGTAGTTTGGGGAAGTCAAAATGCGCATTTTATGTATGTGCAATTGCGTA	15197
OY	3216	tcaatgggcgttggagaagaataatgttggagcccttgaanaagttgaagggccaactct	3275
Db	15198	TCAATGGGCGTGGGAACGAAATATATGTGTGACCTCTGAAAAGTTGTAAAGGGCCCAATTTCT	15257
OY	3276	aagatcttcttaagggagcaagaattaaatgtgtgttagcgtcgaagctatagtttgttggaa	3335
Db	15258	AAGTATTTCTTACGGCAGCCAGCAAGTAAATGTGTGTGACGCTGAGAGTATGTGTTGTGGAA	15317
OY	3336	cgagagcgcaatttttttttlltaacatgtgaacaatgaaaccaacaacaacatlttataaa	3395
Db	15318	CGAGGCGCAATTTTTTTTTTTTTAAACATGGAACACATGAACCAACAAACATTTTAAAA	15377
OY	3396	ctaaaatgataaatttgttaaatagtttcttgactttaaactttaaagtgtttcttgaagt	3455
Db	15378	TTTAAATGCAATATTTTGTAAATATGTTTTTACCTTTTAAAAATTTTAAAGCTCTTTTGTGAGTGT	15437
OY	3456	gaaaagcttgagtaaaactatttgcacatgtgttctcagaanaaggaagaagaacaacaag	3515

Db	15438	GAANAAGTACGAAAACCAATTTGACCTGGTTTTCAGAAAAGCAAAAAGAACCAACCAAG	15497
QY	3516	gaatlgaaacagcgagggagactcctaaccatllcatcattctgcnaatgctactg	3575
Db	15498	GAATTTAAACACAGCGAGGAGATCTTTAATCCCAATTTTATCATTTTCAG-AAAATGTACTG	15556
QY	3576	ttttagaagtgatcacaatlcgaatlgaaatctcttgaaatcgttgtaacatccgact	3635
Db	15557	TTTTAAATGATTTTACATATATACATGTGAATTTCTTGAATCCTGTACAAATCTGCACAT	15616
QY	3636	gatacacaatbtaaatlaattgctgttcgcgaattgacaaatccacacccaatgggga	3695
Db	15617	GTTTAAATCATGTAAATTAATGTTGTGTGCTGATTTGGCCAAATCTCACCAACCAATGGGGA	15676
QY	3696	ggtatcatcagtttgaagaagac-gtgttaacccggaatltgaattgtctcgtgatgtgtgaac	3754
Db	15677	GGATATACATGTTTGAAGAACCTGTATACCTAGTAATTTATTTGTTCTTGATGTTGTACTC	15736
QY	3755	aataagaagtgctttggagaagaagacatgctgtgtgagacagtgctctctctttgycga	3814
Db	15737	AATAAGAGTGTTTGGAGGAGACATGTGTGTGAGACAGTGCTGTTCTTTTGTGTCCAG	15796
QY	3815	cctctgtatgatgtttgtlaagaacacgtcttgtaagaacatlgaaataatgtcgtctttggcc	3874
Db	15797	CTCTGTGATGATGTTTGTGAAGACCATGTTTGTGAAGACATGAATTAATTTGCTTTTGGCC	15856
QY	3875	aa 3876	
Db	15857	AA 15858	

RESULT	5
LOCUS	AC087507
DEFINITION	Homo sapiens chromosome 18 clone RP11-748120 map 18, WORKING DRAFT
ACCESSION	AC087507.3
VERSION	GI:13940683
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Plimates; Catarrhini; Hominiidae; Homo.
AUTHORS	Birren,B., Linton,L., Nisbaum,C. and Landier,E.
JOURNAL	Homo sapiens chromosome 18, clone RP11-748120
REFERENCE	Unpublished 2 (bases 1 to 192022)
AUTHORS	Birren,B., Linton,L., Nisbaum,C., Landier,E., Allen,N., Anderson,S. Barra,N., Bastien,J., Boguslavsky,I., Boukhgalter,B., Brown,A., Calmarete,A., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., Dearlallano,K., Dewar,K., Diaz,J.S., Dodgson,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardina,S., Glinde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagob,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karates,A., LaRocque,K., Lamazares,R., Landers,T., Lechickzy,J., Levine,R., Liu,G., Maclean,C., MacDonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Menous,T., Mihova,T., Mienna,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severly,P., Soungue,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausen,N., Suramanian,A., Talamas,J., Testeys,S., Theodore,J., Travers,N., Trevis,N., Trigglio,J., Vassiliou,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zahoun,J., Zemsek,L., Zimmer,A. and Zody,M.
TITLE	Direct Submission
JOURNAL	Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA



## COMMENT

On May 4, 2001 this sequence version replaced g1:12957877.  
 All reads were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L12322

Center clone name: 748-I-20

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 186183 bases at least Q40

Consensus quality: 190373 bases at least Q30

Insert size: 194000; agarose-fp

Insert size: 190922; sum-of-contigs

Quality coverage: 6.1 in Q20 bases; agarose-fp

Quality coverage: 6.2 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1      6805: contig of 6805 bp in length
*      6806 6905: gap of 100 bp
*      6906 8180: contig of 1275 bp in length
*      8181 8280: gap of 100 bp
*      8281 9283: contig of 1003 bp in length
*      9284 9383: gap of 100 bp
*      9384 11168: contig of 1785 bp in length
*      1169 11268: gap of 100 bp
*      11269 13778: contig of 2510 bp in length
*      1379 13878: gap of 100 bp
*      13879 28641: contig of 14763 bp in length
*      28642 28741: gap of 100 bp
*      28742 43445: contig of 14704 bp in length
*      43446 43545: gap of 100 bp
*      43546 62579: contig of 19034 bp in length
*      62580 62679: gap of 100 bp
*      62680 86613: contig of 23934 bp in length
*      86614 86713: gap of 100 bp
*      86714 125447: contig of 38734 bp in length
*      125448 125547: gap of 100 bp
*      125548 157506: contig of 31959 bp in length
*      157507 157606: gap of 100 bp
*      157607 192022: contig of 34416 bp in length.
```

## FEATURES

## SOURCE

```

1. 192022
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="18"
  /map="18"
  /clone="RP11-748120"
  /clone_lib="RPCT-11 Human Male BAC"
  1. 6805
  /note="assembly_fragment"
  clone_end:SP6
  vector_side:left"
  misc_feature
    6906..8180
  /note="assembly_fragment"
    8281..9283
  misc_feature
    /note="assembly_fragment"
    9384..11168
  /note="assembly_fragment"
```

```

misc_feature      11269..13778
                  /note="assembly_fragment"
misc_feature      13879..28641
                  /note="assembly_fragment"
misc_feature      28742..43445
                  /note="assembly_fragment"
misc_feature      43546..62579
                  /note="assembly_fragment"
misc_feature      62680..86613
                  /note="assembly_fragment"
misc_feature      86714..125447
                  /note="assembly_fragment"
misc_feature      125548..157506
                  /note="assembly_fragment"
misc_feature      157607..192022
                  /note="assembly_fragment"
                  clone_end:77
                  vector_side:right"
BASE COUNT      60131 a 34828 c 35147 g 60815 t 1101 others
ORIGIN
```

```

Query Match      69.9%; Score 2720.4; DB 2; Length 192022;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2749; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

QY 1116 ttcagtcctctgttaagtgacctgtacatgcgcaaaagaagatccagaaga 1175
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187209 TGAGATCCCTATGTAAGTGAACCTGTACCATGCAAAAAGAAATCCCAAGAGAA 187268

QY 1176 gactcatggaagaatgcaaccccaatcagtgctcaatgaagctgttcttgatat 1235
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187269 GACTCATGGAAGAATGCAACCCCAATGCAAGTTCATGACCTGTGCTTTGATAT 187328

QY 1236 tccctgtgagggccttgaagataaagtgatgaaatcttggcttgatcgaagg 1295
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187329 TCCTGTGAGGGCCTTGAAGATTAAGTGTGAATTTGGTTTGATTCGAAAGGGG 187388

QY 1296 gtcccgaaatgaagatcctggcagcttgatcttggtgacgacgagaagaaagctgtg 1355
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187389 GTCCCGAAATGAGGATATCGGCGAGTTAGCTTGGGTGCGACGCAAGAACTGCTG 187448

QY 1356 agagcactggaagaagatcgttgactaccaccaagagaacaatcgcaagtgcaagtgct 1415
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187449 AGAGCACTGGAAGAAGATCTGTACTACCCCAAGACAAATTCGCAAGTGCAGCTGCT 187508

QY 1416 ctgtgagtgtagcatcctagccgtgagtgagtaagctttaaaggttttactaggcaagg 1475
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187509 CTGTGATGTTAGCATCTAGCCGCTGAGTTGGAATTAAGCTTTTACTAGCAAGAG 187568

QY 1476 aaattctcttcttctatatttgatctgcaagcttgggaaatcaagctacctcttctgtg 1535
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187569 AAATTTCTCTCTCTTCTTAATTTGAGATTGCAAGCTTGGAATCAACCTCTTTTGTGG 187628

QY 1536 ttctgttgcttgtagaagaatgagttgaattagtagaccagaagaaagtaactcaaatgtga 1595
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187629 TTCTTTGTTGCTTACGAATGGAATTAATTTAGTAGCCCAAGAACTTCAATTTGTGA 187688

QY 1596 ttatgataattccctattatagaagagtgtagataaatttcataagaatattcaatat 1655
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187689 TTATGATTAATTTCCCTATTATTATAGAAGAGTTGGAATTAATTTTCATTAAGATTAATTA 187748

QY 1656 ctccctcagaattacagagtgatataactaagaagatagtcagacatttataaactgtgcc 1715
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187749 CTCCCTCAGATTACCAAGTATTAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 187808

QY 1716 agaattccaattataatgtagaactatcattggaagacgtgcacaaaagtaagtaagtgta 1775
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187809 AGAATCCCAATTAATTAATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 187868

QY 1776 ttaagaatttaaaacgaaagtaagcttgcttgcttgtagaaattatcatcattatcttcag 1835
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 187869 TTAAAGATTTAAAAACCAAAAGTATGCTTGCTTGGAATTAATTAATTAATTAATTAATTAATTA 187928
```

Oy	1836	gttgggggaatcaatcttcttcttaaccaaagatctaaaaaaatgctccccagttctg	1895
Db	187929	gttgggaattacatttttttctttaattccaaagatctaaaaaattgctccctcagtttgt	187988
Oy	1896	attatatactctgcatactgcacaatggtgtgcccgcataaaagatctgttcattcc	1955
Db	187989	attttattaatctgcctatgtgcgaattgggtgtccctgcataaaaaagattctgtcatttcc	188048
Oy	1956	agtttggttgtaataatttctgactgaattatcatataagaagtaactcagattctatcaa	2015
Db	188049	agtttggtttgttaattttatttttgatgcgaattttattcatnaaagtaactcagatttctcaa	188108
Oy	2016	aagagacatgaacaacacatgcagagaatacttctatcaaaaggctgagttctagaacactg	2075
Db	188109	aagagacatgcgtacacagctgcagaaattttttttatcaaaaaggctgcagttgagaacctgtgc	188168
Oy	2076	ctgaaatataatcttctcccccctaaaggttacaatgtagtgcacaaattctgtaaaata	2135
Db	188169	ctgaaatattattttttttctcccccctaaaggtttacatgtgactcaaaatttttgtaaaatata	188228
Oy	2136	acctcaataaagacaatgagcccttgattatcaatgcagccctgcacaagccctcaatgtagc	2195
Db	188229	acctcaatataagacaatgagcccttgattatcaatgcagccctgcacagccctcaatgtagc	188288
Oy	2196	ctgagaaatcccatcatgctacctcttgagaattgtctgtaattagttagttgataaagaataa	2255
Db	188289	ctgagaaatcccatcatgctacctcttgagaattgtctgtaattagttagttgataaagaataa	188348
Oy	2256	acttcaacatgaatacagaatgtagaagatgcgaattctctctatagaagaatgataatgtgc	2315
Db	188349	acttcaacatgataatccagattatgaagaatgcgaattttctttatagaagaatgataatgtgc	188408
Oy	2316	aagctgtaacttttaagccatctgtttgtacccaagaatgcagatgcacactgaactcat	2375
Db	188409	aagctgtaacttttaagccatctgtttgtatcccaagaatgcagatgcacactgaactcat	188468
Oy	2376	ttaattattgtccccccgaaaaagaataagaatgtaacttgaaaagacgtgagaactttc	2435
Db	188469	ttaattattgtttgtccccccgaaaaagtttaagatgtatctttaaanaagcctgtgaagattttt	188528
Oy	2436	tacattgccaagataaagaatgcttaacttaacccaacaaatgtaagactacaataatcgt	2495
Db	188529	tacattgccaagataaagaatgcttaacttaacccaacaaatgtaagactacaataatcgt	188588
Oy	2496	tcaagagcaattctaatataaatcttaacatgcttcaatgcgcaaaatgatgctgagctgtcaaa	2555
Db	188589	tcaagagcaattctaatataaatcttaacatgcttcaatgcgcaaaatgatgctgagctgtcaaa	188648
Oy	2556	ttagcacaacaagaatgctgtctcaacatctcttctcagactcaattgtcttgaagctgtg	2615
Db	188649	ttagcacaacaagaatgctgtgttcaacatctcttctcagactcaattgtcttgaagctgtg	188708
Oy	2616	tcaatagagcaatcttaacagactgtgctctgttctatcttctcaagatgcagagttcctaaa	2675
Db	188709	tcaatagagcaatcttaacagactgtgctctgttctatcttctcaagatgcagagttcctaaa	188768
Oy	2676	ttaactcagaacactgtgaataaagcctgcagaccctgtgattcttggaaaagaatagactc	2735
Db	188769	ttaactcagaacactgtgtgatttaaaagcctgcagaccctgtgatttttgamaaagaattagactc	188828
Oy	2736	gagaagaatgaatcaactatttggagttctcttagagaagatagaatggaacttgaagaactc	2795
Db	188829	gagaagaatgaatcaactatttggagttctcttagagaagatagaatggaacttgaagaactc	188888
Oy	2796	gaattatataataatgcagaatgtagaatlaagctcactgaaaaaaattacatttgaatcag	2855
Db	188889	gaattatttttaaaattatcacaagttgaaaatttaagctcactgaaaaaattttacattttgaatcag	188948
Oy	2856	gttttgatgcagtaactttagaagtttttggaaatgctgtttgataatcatcagaatgtttgaaa	2915
Db	188949	gtttttgtgtcactacttttagaagatttttttgaaatgctgtttgattatcacagatgtttgaaa	189008

OY	2316	tttcttgaaaataatgcatctttccaaacacttttaacacgcctttttattatcactacacgcctaatg	2975
Db	189009	ttctatgacaaaataatgcattttttccaaacactatgacatgcgtttttatgactatgcctaatg	189068
OY	2976	ttaagaataatgtatatacattcctgtatgataagaataataaatacaccctcttttctgtgc	3035
Db	189069	ttatgaaaataatgatttacttctgtatgataatgacaaagatttataatataacacctttttgtgtgc	189128
OY	3036	tttaaaaagacttggagatttaaaaagaacatalttcccaacattgtcttcattccactac	3095
Db	189129	tttttataataacttttgaggatttataaaaagcattatttcccaatattgtctttatgtccactac	189188
OY	3096	aaagtcacctcaacagacatctgtcctcaactcggcatctctgtgaaagcaacatgaaatgaa	3155
Db	189189	aaagtcacctcaacacacatctgtgtgtccactcggcatctctgtgaaagcaacatataatgtaa	189248
OY	3156	cctgtagtaagtgtgtagtttttggggaagtcacaaatggtccatttatgtatgtcatgtgta	3215
Db	189249	ctgtatgtactgtgtgtatggtttggggaagtcacaaatggtccatttttatgtatgtcgttgtaa	189308
OY	3216	tcatgtgcccgttggaacacagatalatgtctgagaccttgaaaaagtttgtaaggggccaactc	3275
Db	189309	tcatgtgcccgttggaacacataatgtgtgtgacctcttgaaaaagttgtataggggccaactct	189368
OY	3276	aaglatcttctccagcgacccagaaagtaabgtgtgtagcaactgtaggtatgtgtgttga	3335
Db	189369	aaatattttcttccagcgacccagaaagttatgtgtgtgacctgtaggtatgtgtgttga	189428
OY	3336	cgaagccgactttcttcttcttatacatgtaacacagaaacacaaacacatttttaaa	3395
Db	189429	cgaagcccgatttttttttttttttttatacatgtaacacataaaacaaacatttttttaaa	189488
OY	3396	ttaaaaatggataaattctgtlaaatagttcttagctctttaaatttaaaagtgtttttgagt	3455
Db	189489	tttttaataatggatataattgttataatgatttttttttttttttttaataatgatttttgagt	189548
OY	3456	gaaaagttcgaglaaaacactattgtgcactcgtgtttccagaaaagagaaaaaacaacaag	3515
Db	189549	gaaaataattgaglaaaacactattgtgcactcgtgtttccagaaaagagaaaaaacaacaag	189608
OY	3516	gaattcgaaaacagcgaggagatccttaatacctaattccatacttctgcgaaaatgtaactg	3575
Db	189609	gaatttttataaacagcgaggagatccttattactccttataatttttatatttttcg-aaaattgactg	189667
OY	3576	ttttagatgtaatttaacaatalcaatgtagataactgttagatcctgttacaactcctgcact	3635
Db	189668	ttttttataaataatttttaacaataatgtaattctgttgagattcctgtttcaaatctctgcact	189727
OY	3636	gatttaacaatgtaataatthaatttgttctgcgtattgccaactcaccacccaatgaggaa	3695
Db	189728	gatttttaaacatatttataatttaattgattgtgtgattgtggccaatctcaccacccaatgggaa	189787
OY	3696	ggtatacaatgatttggaagac-gttagtaacctcgtaattgatttgttctgtaagtgttaactc	3754
Db	189788	ggattttacatgatttggaagacggtgtgactacgatttaattgattgtttcgatgattgttaactc	189847
OY	3755	aaataagaagtgttttggaaagaaagcattggtgttgagaacagtgtcgtctcttctgtgcag	3814
Db	189848	aaatagaaagcgtttttggaaagaaagcattggtgtgtgagacacagctcgtctcttctgtgcag	189907
OY	3815	ctctgatatgatttctgttaagacacatttctgttaagacacgaaataaatgctgcacttgcgc	3874
Db	189908	ctctgtgatatgatttctgttaagacacatttctgttaagacacgaaataaatgctgccttttgccc	189967
OY	3875	aa 3876	
Db	189968	aa 189969	

LOCUS	AC091039	159135 bp	DNA	HTG	11-AUG-2001
DEFINITION	Homo sapiens chromosome 18 clone RP11-40345 map 18.*** SEQUENCINGS				
RESULT	6				
AC091039/c					



OY 2136 acctacataagaacatgagccttgatattcaatgcctgtcaagaagctcaagtcagtcg 2195  
 |||||  
 Db 146033 ACCACATATAAGAACCATGCGCTTGATATATCACTGCTGTCAAGACCTCAGTGGCG 145974  
 OY 2196 ctgagaacatcccatgaccttgatgaatgctgaatgaatgctgaatgaatgaatga 2255  
 |||||  
 Db 145973 CTGAGAAATCCCTATGACCTTGTGAAATGTGTGAATTAAGTAAAGAAATTA 145914  
 OY 2256 acctcaactgaatccatgctgaatgctgaatgctgaatgctgaatgctgaatgctg 2315  
 |||||  
 Db 145913 ACTTCAACTGAAATCCAGTGAAGTGAATTTCTTATGAAATAGGTATAGTGTGC 145854  
 OY 2316 aagtgactcttgaagcagtcgttgtaacagagtgagcagcagcagcagcagtcag 2375  
 |||||  
 Db 145853 AAGTGACTTTTAAGGCCATCGTTGTACCCAGACTGGCCACCTAAGTCTTCAT 145794  
 OY 2376 ttaattatgtcccccagaaagatgaagtgctacttgaaagactgtgaagattttt 2435  
 |||||  
 Db 145793 TTAATTTATTTGTCCCGAAGAAAGATTAAGATGCTACTTGAAGAGCTGTGAAGATTTT 145734  
 OY 2436 taactgtccagataaagaatgttacttaacacaacaatgtaagaatcaaatcgt 2495  
 |||||  
 Db 145733 TCAATTCGCGATTAAGAAAGTGTACTTAACCAACAAATGTAAGACTTAACAAATCGT 145674  
 OY 2496 tcaagagaactcttaataataattacataatgttcaagcaaaatagcttaagctgtcaaa 2555  
 |||||  
 Db 145673 TCAAGACCAATTCATATATTAATTTACATATGTTCCAGCAAAATATGCTTAGCTGTCAAA 145614  
 OY 2556 ttaagacaacaagaatglttcaactatcttctcaagcttaatttgcttgagctgttg 2615  
 |||||  
 Db 145613 TTACACACAAAGAAATGAGTGTCACTATCTTTCTAGCAATATTTCTGTGAGCTGTGG 145554  
 OY 2616 tctatagaagcagttcaagaactgtgtctgtatcatcttccaatgcaaggtctga 2675  
 |||||  
 Db 145553 TCTATTAAGACAGATTTGACAACTGTGTCTGTATCATTTTCCATGCGCAAGGTCTGA 145494  
 OY 2676 tctcatagaagaactgttaagcttaagaactgtcaacccgtgtaattatctgaagaatga 2735  
 |||||  
 Db 145493 TTCAATTCAGAACTGTATTAATTAAGCTGCACCCGTATTAATTTGAAAGATTTAGCTT 145314  
 OY 2736 gaagtgatctcaactatctgagtcctctgaagaatgctgtgaactgaactga 2795  
 |||||  
 Db 145433 GAGAGTAATGTCACTATATTTGAGTCTTAGAGAGATGAGTGAAGTGAAGTGAAGT 145374  
 OY 2796 gaattatataatgcaagctgaagaatgaagctcaactgaagaatctcaattgagtcag 2855  
 |||||  
 Db 145373 GAATTTATTAATATGCAAGTTGAAATTAAGTCTACGTGAAATTTACATTTTGAAGTCA 145314  
 OY 2856 gttttgtgtaagactgttagcagtttttgaagaatgctgtgaatgaactgtgttga 2915  
 |||||  
 Db 145313 GTTTTGTGCTAGTACTTTGACAGTTTTCAGAAATGCTGTTTATATCACTGTTGTAA 145254  
 OY 2916 ttcatataaagaactgttccaagaactataagaactgttcatgaactgaactga 2975  
 |||||  
 Db 145253 TTCTATCAAAATGCACTTTTCCAAACATTAATACATCTTTTATACATGCTCAATAG 145194  
 OY 2976 taaagaagaatgattacattctgtatgtaacaagaatgaagaatcaactcttctgtgc 3035  
 |||||  
 Db 145193 TAAAGAAATATGATTAATCACTTGTATGTAACAAAGTAAATTAATTAACCTTTTGTGTC 145134  
 OY 3036 tttaaatgactgttgatataaagaatataatccaactatgttctcatcactgc 3095  
 |||||  
 Db 145133 TTTAAATATGCTTTGGGATTTAAAGATATTTTCCCAATCATTTGCTTCATTCACCTAC 145074  
 OY 3096 aaagtcacccacacagatctgtccactcgcagcactcctgtgaaagaacataaataga 3155  
 |||||  
 Db 145073 AAAGTCACTCCACAGCATCTGCTCCACTCGGCATCTCTGTGAAGCAACATGAATGAA 145014  
 OY 3156 ctgtaagaagtgtagtctgttggaagtcacaatgagcctttaaagtagtgcactgtga 3215  
 |||||  
 Db 145013 CTGAGTAGGTGTGTAGTTGGGAAAGTCAATAGCCATTTTATGATATGTCATTTGGTA 144954

OY 3216 tcatggccgtgagacagaatataatgttgagacctctgaagaatgtgaaagggccaactc 3275  
 |||||  
 Db 144953 TCATGGCCCGTGAGACAGATATATGTGGACGCTGTGAAGAGTGTGAAGGGCCAAATCT 144894  
 OY 3276 aagtatcttcacgcgcagcagaaagtaagtgtgtgagagcgcgcgcgcgcgcgcgcgc 3335  
 |||||  
 Db 144893 AAGTATCTTCACGCGCACCCAGAAATTAATGTGTGTGAAGTGAAGTGTGTGTGTGA 144834  
 OY 3336 cgaagccga-tttttttttaaactgaagaacataaagaacacacacacacacacacac 3394  
 |||||  
 Db 144833 CAGGCGCAATTTTTTTTTTTTTTAAATGATGAAACATGAACCAACCAACCAATTTTAA 144774  
 OY 3395 attaaatggaataatgttaaatgttttaagcttttaaatgaatgaatgaatgaatga 3454  
 |||||  
 Db 144773 ATTAATATGATTAATTTTAAATAGTTTAACTTTTAAATTTAAAGTGTGTGAATG 144714  
 OY 3455 tgaagaatgtgaataactatctgcaactgtgttcaagaagaagaagaagaacaa 3514  
 |||||  
 Db 144713 TGAAGATGTGAATTAATCTATTTTCAACTGTGTTTCAAGAAAGAAAGAAACAA 144654  
 OY 3515 ggaattgaacagcagggagacatcttaatacctaattcaactctgtgcaaatgta 3574  
 |||||  
 Db 144653 GGAATTTGAAGACGCGAGAGATCTTAATACCTTAATTTCACTATTTCTG-AAATGTACT 144595  
 OY 3575 gtttgaagatgattacaataatcaatgtgaatatcttgaatccctgttacaactcctgcac 3634  
 |||||  
 Db 144594 GTTTTGAATGTATTACAAATATCAATGTGAATATCTTGAATCTGTACAAATCTGTGAC 144535  
 OY 3635 tglataacatgtaataatgattgtgtgtgtaagcgaactcaccacacacacacacacac 3694  
 |||||  
 Db 144534 TGTATTAAACATGTAATTAATTTGTTGTCTGATTAGCCAAATCTCCACCAAAATGGGG 144475  
 OY 3695 aggtatcacatglttgaagaac-gtgaactcgcgtgaatgattgtgtcgtatgttaact 3735  
 |||||  
 Db 144474 AGGTATACATGTTTGAAAGACATGTAACACATAATGATTTGTTCTGATGTGTACT 144415  
 OY 3754 caatagaatglttgaagaagaacatgctgtgtgaagaacagctgtgtgtgtgtgtgtgtgt 3813  
 |||||  
 Db 144414 CAATGAAGTGTGTTGGAAGACATGATGTGTGAAGACAGTGTGTGTGTGTGTGTGTGTGT 144355  
 OY 3814 gctcgtatgattgttgaagaacatglttgaagaacatgattgtgtgtgtgtgtgtgtgtgt 3873  
 |||||  
 Db 144354 GCTGTGTATGATGTTTGAAGACATGTTGTGAAGACATGAATTAATTTGCTGTGTGTGTGT 144295  
 OY 3874 caa 3876  
 |||||  
 Db 144294 CAA 144292

RESULT 7  
 AF299075  
 LOCUS Homo sapiens synapcotagmin IV mRNA, complete cds. 04-DEC-2000  
 DEFINITION AF299075  
 ACCESSION AF299075.1 GI:11528336  
 VERSION  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Ferguson, G.D., Chen, X.-N., Korenberg, J.R., and Herschman, H.R.  
 TITLE The Human Synapcotagmin IV Gene Defines an Evolutionary Break Point  
 between Syntenic Mouse and Human Chromosome Regions but Retains  
 Ligand Inducibility and Tissue Specificity  
 JOURNAL J. Biol. Chem. 275 (47), 36920-36926 (2000)  
 PUBMED 10938284  
 REFERENCE 2 (bases 1 to 1671)  
 AUTHORS Ferguson, G.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-AUG-2000) Pharmacology, University of Washington,  
 Health Sciences J681, Seattle, WA 98195, USA  
 FEATURES Location/Qualifiers

```
source
1.1671
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12.3"
94..1371
/codon_start=1
/product="synaptotagmin IV"
/protein_id="AAG37229.1"
/db_xref="GI:11528337"
/translation="MAP1TTSKEPDELPYVVGIFSAEGLVETVLSLFAMICQKRSK
SKRTPPYKFEVHLKGVLDIYPENLNSKKRFGADKNEKRPVAPKNSLHLDLEKBDLN
GNFPTNLKPGSPDLNENATPFLFEGESYSPESLSISLSEKQEKLEKTLFES
LEYNEKRAFYVNIKEARGLPAMDOSMTSDYIKMTILPEKHVYKTPVLRKTLDP
FDETFPIGYTOIOLAFHETLISPDRESBDIIGVLIPLSGIELSEKMLNRE
IKRNRKSSSGGELLISLCYSTNTNTLVVYLRKRLPKSPVSGISDPYVYNLYHA
KKRISKKTTHVKCTPNVFNELFVFDIPCEGLDIEIVSEVLDSERSRNEIQLV
LGAALGGEGHEMKEDICDYPRIQIAKMHVLCDG"
BASE COUNT      513 a      330 c      372 g      456 t
ORIGIN

Query Match      42.4%: Score 1651; DB 9; Length 1671;
Best Local Similarity 99.9%: Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 67 gaaagccggcgcttgatccaggaagtgaatccagccagcgaggtttccc-ttcagca 125
DB 9 GAAAGCCGGCGCTTGATCCAGGAAAGTAAATCCAGCCAGCGAGTTTCCCTTCAGCA 68
QY 126 cctcgcagacaacacgcagtaaaaaatgctccgatcacaccacccagcggaagaattga 185
DB 69 CCTCGGACAGAACACGCAGTAAAAAATGGCTCGATCACACCAGCGCGGAGAAATTGA 128
QY 186 tgaatcccccacagtggtgggagcttccagtgcaattgagcctggtcttccagctctct 245
DB 129 TGAATCCCCACAGTGGTGGAGTCTTCAGTGCATTGGCGCTGCTTCACAGTCTCTCT 188
QY 246 ctttgatgatctgctgctcagagaagaatcatccaagtccaagaactcttcataca 305
DB 189 CTTTGATGATCTGCTGCTCAGAGAAATCATCAAGTCTAACAAACATCTCTCATACAA 248
QY 306 gtttgatcgttgcttaaggagattgataltaacctgaaacctaataatgacaaaaga 365
DB 249 GTTTGTCATGTCCTTAAGGAGTTGATATTTCCTGAAAACTTAATGCAAAAAGAA 308
QY 366 gtttgagcagatgataaaaaatgaagtaagaataagcagctggtgccaaagaattca 425
DB 309 GTTTGGAGCAGATGATMAAAATGNAAGTAAAGATAAGCCAGCTGTGCCAAAGATTCA 368
QY 426 gcatcgtgattctgaaagaagagatctcaatgcaatttccccaaaaccaactcaacc 485
DB 369 GCATCTGATCTTGAAAGAGAGATCTCAATGGCAATTTTCCCAAAACCAACTCAAAAC 428
QY 486 tggcagcttcctctgcatctggaagatgcaaccccggaagcttcttlttgaaggagaaaaga 545
DB 429 TGGCAGTCCCTTGATCTGAGAGATGCAACCCCGAAGCTCTTTTGAAGGGGAAAAAGA 488
QY 546 gtcagattccctcgtgagatttaagtcacagcacttcccttacttcagaagagaacaaga 605
DB 489 GTCAAGTTTCCCTCGAAGATTAAAGTCCAGCACTTCCCTTACTTTCAGAGAGAAACAAGA 548
QY 606 gaagctgggaactcttcttctcctctagaatacaacttcgagaagagaacatttgtgt 665
DB 549 GAAAGCTGGGAACTCTTCTTCTCTCTTAGATTAACAATTCTGGAGAGAAACATTTGGGT 608
QY 666 caatatacgaagaccgctggtcttgcagccatgataagcagctgaatgacatctgaacc 725
DB 609 CAATATCAAGGAAGCCCGTGGCTTGGCAGCATGATGAGCACTGATGACCTCTGACCC 668
QY 726 atatatcaaatgaagatctctccagagaagaagcataaagtgaataactagatgctgag 785
DB 669 ATATATCAAAATGACGATCTCTCCAGAGAAAGACATAAAGTAAAGTAAAGTGAAGTGTGAG 728
```

```
QY 786 aaaaaccttgatccagcttltgatagacctltacatltcatgagataccaccca 845
DB 729 AAAAACCCTTGATCCAGCTTTTGATGAGACCTTTACATTTCTATGGGATACCCTACACCA 788
QY 846 aatccaagaattggccttgcaacttcaaatlttgaglttggacaggttttcaagagatga 905
DB 789 AATCCAAAGAATTGGCCTTGCACTTACAAATTTGAGTTTGAACAGGTTTCAAGAGATGA 848
QY 906 tatcatltggggaagcttctaattccctctcccggaattgaattatctcgaaggaaaatgt 965
DB 849 TATCATTTGGGGAAGTTCTTAATTCCTCTCGGGAATTGAATTTCTGAAGGAAAAATGTT 908
QY 966 aatgaatagagaatcatcaagaagaatgttaagaaagcttcaagagccgaggtgaatct 1025
DB 909 AATGAATAGAGATATATCAAGAGAAATGTTTGAAGAGCTTTCAGAGAGGGGTAGTTACT 968
QY 1026 gatctctctgtcatcagtcacacccaacactcaactgagtgctgtcttaaaagctcg 1085
DB 969 GATCTCTCTGCTATTCAGTCCACCAAAACCTTAAGTGTGTTTAAGCTCG 1028
QY 1086 acaatcgcttaaatcgtatgtgtcgggaacttcaagatccctatgtcaaatgtaacctgta 1145
DB 1029 ACATTCGCCCTAAATCTGATGTGTCCGACTTTCAGATCCCTATGTCAAAGTGAACCTGTA 1088
QY 1146 ccaltgcacaaagaagatctccagaagaagaactcatgtgaagaagaatgcccccaatgc 1205
DB 1089 CCAATGCCAAAGAAAGATCTCCAGAGAGAGACTCATGTGAAGAAATGACCCCCCAATGC 1148
QY 1206 agtctcaatgaagctgttcttcttatacttctgtgagggccttgaagataataagtg 1265
DB 1149 AGTGTCAATAGACCTGTTGTTGATATTCTTGATGAGGCTTGAAGATATAAGTGT 1208
QY 1266 tgaatttttggttttgatcttgaagggggtccgaaatgagtgatcggcgagttagt 1325
DB 1209 TGAATTTTGGTTTGGATTCTGAAAGGGGGTCCGAAAGAGATGATCGGCAAGTTAGT 1268
QY 1326 ctgggttgacagcagcagaagaacttggtgagaagacactggaagaagatctgttactacc 1385
DB 1269 CTTGGGTGACGACGACAGAGAACTGTGTGAGAGACTGCAAAAGAGATCTGTACTACCC 1328
QY 1386 caggagaacaattgccaagtgcacgtgctctgtgaatgtaagcatalcctacgttgagt 1445
DB 1329 CAGGAGACAAATTTGCCAAGTGGCAGTGTCTGTGATGTTTACATCTTACCGCTGAGT 1388
QY 1446 ggaacttaaggttcttactaaggcagaaggaatlttcttctctatataatgtatgca 1505
DB 1389 GGAACCTTAAAGGTTTACTAGGCAAGAGAAATTTCTTCTTATATTGGAATTGCA 1448
QY 1506 agcttgggaatcaagctacacttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1565
DB 1449 AGCTTGGGAATCAAGCTACCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1508
QY 1566 agtagacccaagaatgaacttcaaatgtgtatataatgtaatttccctatattatagaag 1625
DB 1509 AGTAGACCAAGAAAGTAACTTCAAAATGTATATGATTAATTTCCATATTATTAGAAAG 1568
QY 1626 ttggataaatlttcaataagatatataatcttccagatctacagtgatataactagg 1685
DB 1569 TTGGATTAATTTTCAATGAAGTATTAATCAATCTCTTCAAGTATTCAGATGATTAAC 1628
QY 1686 aatagtcagacatttatgaatactgtgcagaaatcccaaat 1728
DB 1629 AATAGTCAGACATTTTATGATACTGTGTGCCAGAAATCCCAAAATT 1671
```

```
RESULT 8
AC068125/c AC068125 170218 bp DNA HTG 24-AUG-2000
LOCUS Homo sapiens chromosome 18 clone RP11-403A5, WORKING DRAFT
DEFINITION
SEQUENCE, 16 unordered pieces.
ACCESSION AC068125
VERSION AC068125.5 GI:9887816
```



Db 155707 ACATTGCCAGATAAAAAGTGTACTTAACCAACAAACAAATGTAAGCTACCAAAATCGTT 155648  
 QY 2497 caagagcaattcctaataataatttacaatgcttcacgcaaaatatgcttgctgcaaat 2556  
 Db 155647 CAAGAGCAATTCCTAAATATATATTTACATATGTTCCACGCAAAATATGCTTAGGCGTCAAAAT 155588  
 QY 2557 tagcacaacaagaatgctgcttccacatcttctcagcctaatttctcagcgtctg 2616  
 Db 155587 TAGCACACAAAGAAATGTTTTCACATCTTTTCTTAGGCTAATTTTCTTAGGAGCTTTGT 155528  
 QY 2617 ctatagagcaagtttccacagctgctgctgtaacatcttccagtgccaggtctgtaaat 2676  
 Db 155522 CTATAGAGCAGGTTTACAGACTGTGTCTGTATCATTTTCCAGTCCAGGCGTTCTGAAAT 155468  
 QY 2677 tcaatcagaacctgctgtaagttaagctgcacccctgctgatatcttgaagaagaattgacttg 2736  
 Db 155467 TCATTTCAGAACCTGTTAGATTTAAAGCTGCACCCCTGATTTATTTGAAAGAAATTTACCTTG 155408  
 QY 2737 agagtaatgctcaataattttagcttcttagagaagatgagtggaaccttgagtaagctg 2796  
 Db 155407 AGAGTAATCTCATTATTTAGTCTTTAGAGAGATGAGTGAAGCTTGAGTACAGTGTG 155348  
 QY 2797 aattatlaaatacgcaagttagaatlaagctcactgtaaaatlaacatcttgaagtcag 2856  
 Db 155347 AATTATTAATATGCAAGTTAGAAATTAAGTCTACGAAAAATTTACATTTTGAGTCAGG 155288  
 QY 2857 ttcttgctgaactttagcaatctttaggaagtgtgttgatatacagactgtttgtaaat 2916  
 Db 155287 TTTTGCTAGTACCTTTTACAGAGTTTGAAGATGCTTTGATATACAGGTTTGTAAAT 155228  
 QY 2917 tctatgaaaaatgcatcttccacaacactatacatgcttcttatacactatgcttaatg 2976  
 Db 155227 TCTATGAAAAATGCAATTTCCAAACAACTTATNCATGCTTTTATGATGCTATG 155168  
 QY 2977 aagaagaatgattacatctctgtaatgtaacaagaatlaaacacacctcttttctgct 3036  
 Db 155167 AAGAAAAATGATATCATCTCTGATGTACAAAGATTAATAAATCAACCTCTTTTCTGTCT 155108  
 QY 3037 ttaaaatgacttgggattaaaaaagcatattcccaatcatgcttcttcaatccatac 3096  
 Db 155107 TTTAAATGACTTTGGGATTTAAAAAGCATATTTCCCAATCATTTGCTTCACTCCATAC 155048  
 QY 3097 aagtcacctcacagcaactctgctccacactcgcaactctctgtgaagaacaataatgaac 3156  
 Db 155047 AAGTACCTCACAAGCATTTGCTCCACCTCGGACATCTCTGTAAGAACACATGTAATGAAC 154988  
 QY 3157 tctagtagtctgtagcttgggggaagtaaaatgagcaatcttctgtagctgtagct 3216  
 Db 154987 TGTAGTAGGTGTGTACTTTGGGCAACTCAAAATGGCCATTTTATGTATGTCATTTGTAT 154928  
 QY 3217 catgggcgttggaacagaaatatgttgagacctctgaagaagtgtgaaggccaaacta 3276  
 Db 154927 CATGGCCGTGGACAAATATATGTGGACCTCTGAAAAGTTGTAAAGGGCCAAATTTCTA 154868  
 QY 3277 agtatctctcacgcaagcaagaagttaatgctgtagcagctgaggtalatgcttctgagc 3336  
 Db 154867 AGTATCTTCACGGCAGCAGCAAGTAATGCTGTAGCAGCTAGGCTATGCTTTGGAC 154808  
 QY 3337 gaagcgca-ttttttttttcaacatgaaacaaatgaaccaacaacaacatctttaa 3395  
 Db 154807 GAGGCCGATTTTTTTTTTTTTTAAATGTAAGTAATGAACCAACCAACCAATTTTAAAA 154748  
 QY 3396 ttaaaatgataaattgttaaatagctttagactttaaagtgtttttagagtg 3455  
 Db 154747 TTTAAATGGATATTTGTAAATATGTTTATGCTTTTAAATTTTAAAGTGTTTTGAAGTG 154688  
 QY 3456 gaaaatgtagaataaacttttgcact-gtttccagaagaagaagaagaacaacaa 3514  
 Db 154687 GAAAATTTGGTAAACTATTTTSCAACTGGGTTTCAGAAAAGCAAGAAACAAACAAA 154628  
 QY 3515 ggaatgtgaacagcaggaagatcttaatacctaattcaatctctgcgaatgtagct 3574  
 Db 154627 GGAATTTGAAACAGGACGAGATCTTAATACCAATTTTCATCATTTCTG-AAATGTACT 154569

QY 3575 gttttagaagttattacaataatgtaatatcttgaatctctgcttacaatccctgac 3634  
 Db 154568 GTTTTAGAATGTAATTAACAATATCAATGTAATGTAATCTGTATGTAATCTGTGAC 154509  
 QY 3635 tctatlaaacatgtaataatattgctgtctgtagtaagcaatctccacccaatgag 3694  
 Db 154508 TGTATTAACATGTAATTAATTTGCTGTATGTAATGCAATCTCCACCAACCAATGGGG 154449  
 QY 3695 aggtatcatgcttgaagaac-gtgaactcggtaattgtaattgctgtagctgtagct 3753  
 Db 154448 AGGTATACATGTTTGAAGCACTGTGTACTAGTAATGATTTGTTCTGATGTTGTAAC 154389  
 QY 3754 caatagaagtgttttggagaagcagtggtgtagaagagtgctgtcttcttgcca 3813  
 Db 154388 CATATGAAGCTTTTGTGAAGAACATGCTGTGTGACAGTGTCTGTTTGTGCGCA 154329  
 QY 3814 gctctgtagctgttctgtaagaccatgcttctgtaagcaatgaataatgctgcttgc 3873  
 Db 154328 GCTCTGATGAGTGTGTGAAGCCATGTTGTGAAGCATGAATTAATTCCTGCTTTGCGC 154269  
 QY 3874 caa 3876  
 Db 154268 CAA 154266

RESULT 9  
 MM010355 3992 bp mRNA ROD 23-AUG-1994  
 LOCUS Mus musculus BALB/c synaptotagmin 4 (syta) mRNA, complete cds.  
 DEFINITION U10355  
 ACCESSION U10355  
 VERSION U10355.1 GI:499125  
 KEYWORDS  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 REFERENCE 1 (bases 1 to 3992)  
 AUTHORS Hilbush,B.S. and Morgan,J.I.  
 TITLE A third synaptotagmin gene, syt3, in the mouse  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91, 8195-8199 (1994)  
 MEDLINE 9436712  
 REFERENCE 2 (sites)  
 AUTHORS Perin,M.S., Fried,V.A., Mignery,G.A., Jahn,R. and Sudhof,T.C.  
 TITLE Phospholipid binding by a synaptic vesicle protein homologous to  
 the regulatory region of protein kinase C  
 JOURNAL Nature 345 (6272), 260-263 (1990)  
 MEDLINE 90238548  
 REFERENCE 3 (bases 1 to 3992)  
 AUTHORS Hilbush,B.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JUN-1994) Brian S. Hilbush, Roche Institute of  
 Molecular Biology, 340 Kingsland St., Nutley, NJ 07110, USA  
 FEATURES  
 source location/Qualifiers  
 1..3992  
 /organism="Mus musculus"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /clone="p171"  
 /tissue\_type="Brain"  
 /tissue="syta"  
 /gene="syta"  
 /gene="syta"  
 /codon\_start=1  
 /product="synaptotagmin 4"  
 /protein\_id="AAA20971.1"  
 /db\_xref="GI:499126"  
 /translation="MAPITTSVPEDEIPITVYGISAGVITVSLAMVICORRSAS  
 SKRTPYKRVHLKVDIYIPENLSSQKFGGDKSEVVKRTALPMLSHLDEKRLDN  
 GNPFRANPRAGSSSDLENVTPKLTETKEASPELSSTSLTSEEOEKGLTFLS  
 LEYNEPKAFVNIKEAQLPAMDOSMTSDPYIMKJTLPEKRAHVKKRVLRKTLDPV  
 FGEFTFYGIPIYPHIOELSLHFTVLSFDRFSRDYIVGILVLPISGIELSDGKMLMTRF

IIKRNKSSGSGELIVSLCYOSTTNTLVVILKARHLKSPDVSGLSPYVKNLYHA  
KRISKSKTHVKTPNAVFNELFVDFIPCESELETSVFLVLDSEBSRNRNVLGRVY  
LQATAGSGGSHMKEDCFPRQIAKWHMLDGS"  
2157. .2486  
/gene="sy14"  
/note="C2A repeat region, amino acids 148-257, internal  
repeat homologous to the protein kinase C C2 domain found  
in all members of the synaptotagmin family"  
/citation=[2]  
2556. .2899  
/gene="sy14"  
/note="C2B repeat region, amino acids 281-394, internal  
repeat homologous to the protein kinase C C2 domain found  
in all members of the synaptotagmin family"

BASE COUNT 1079 a 894 c 986 g 1033 t  
ORIGIN  
/citation=[2]

Query Match 33.5%; Score 1305.4; DB 10; Length 3992;  
Best Local Similarity 77.4%; Pred. No. 1.8e-241;  
Matches 1781; Conservative 0; Mismatches 461; Indels 59; Gaps 14;  
QY 115 tcccttcagacctcgagacagacacgcagtaaaaaatgctccgataccaccagcccg 174  
Db 1681 TTTCTTACGACCTCGAGACAGAGACAGCAG-AAAACATGGCTCTTATCCACCACACCCG 1739  
QY 175 gaagaatttgataaaatcccaacagtggtggtggaatcctcaatgacattggtccgtcttc 234  
Db 1740 GTGGAAATTCATGAATATCCACAGTGGGGCATCTCACTGCTTGGCCGTCATCTTC 1799  
QY 235 acaatcctctcttcgacatgacatgctgcttcagagaagaatcatcaagcttaacaagact 294  
Db 1800 ACTGCTCTCTCTTCTGCGATCTGCTGTCAGAGAAATCAGCAAAATCCACAAGACT 1859  
QY 295 cctccatacaagtttgatgacatgcttaagagagtgatattacccttgtaaaacctaat 354  
Db 1860 CCTCCATACAAAGTTTGTGACGCGCTTAAGAGATGTATCTCCAGAAACCTTAAGT 1919  
QY 355 agcaaaaaagagtttgagacagatataaaaaatgaagaagaagaatgaacagcttctgca 414  
Db 1920 AGCCAAAGAGTTTGAGAGATGACAAAGTAAAGGTAAACCGCTCGCC 1979  
QY 415 aagaattcatgacatgacatgacatgacatgacatgacatgacatgacatgacatgacat 474  
Db 1980 AACCTTCCCTGATCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2039  
QY 475 aacctcaaacctggagctccttcgacatgacatgacatgacatgacatgacatgacatgac 534  
Db 2040 AACCCCAACCTGGACGCTTCTGATCTGGAATGACACCCCAAGAGCTTCTTACGGAG 2099  
QY 535 gggggaagaagaagcttccctcagagtttaaaagtcacagactcccttacttaagaa 594  
Db 2100 ACAGAAAGAGGCGCAATTCCTCGAGAGCTTGAAGTCCACCACTTCCCTCAGCTGAGAG 2159  
QY 595 gagaacaagaagaagctgggaactctctcttccttccttaagaatacaactcgaagagaaa 654  
Db 2160 GAGAAACAAAGAGAGCTGGGAGACTTCTTGTCTTACAGTACCACTTCGAGAGAGAA 2219  
QY 655 gcaatttgatcaatacaagaagcccgtagcttcgacagccatgtagtagagagtagatg 714  
Db 2220 GCAATTTGTGTGTAACATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2279  
QY 715 aacctggaaccatatacaaaatgacgaatccctccagagagaagaagcaataagtgaaact 774  
Db 2280 ACCTTGACCCGATCAAAATGACATCTTACACAGAGAGAGAGAGAGAGAGAGAGAGAG 2339  
QY 775 agagctcgaagaagaagccttgatcagacttttgatagacacttcaatcctatgagata 834  
Db 2340 AGAGTGTCTGAGAGAGAGCTGAGACCTGTTTGTGTAACCTTCAATCTATGGAATC 2399  
QY 835 cccatacccaaatccaagaatggccttgacactcaaatgttgagatttgaacaggttt 894

Db 2400 CCTTATCCCCACATCCAGAGAGCTCTCCCTGACACTTTACAGTTCTGATTTTGACAGTTT 2459  
QY 895 tcaagagatgatatcatctgggaagctcttaattccctctccggaattgaattatctga 954  
Db 2460 TCTAGAGATGATGCTATTGGAGAGCTCCTGATTCCTCTTCAAGGATTGAATTATCAAT 2519  
QY 955 ggaagaattgaataatagaagaatatacaagaagaagaattgaagaagctcgaagag 1014  
Db 2520 GGAATAATGTTAATGACATGAGAGATCATCAAGAGAAATCTGAGAAAGCTTCTGCGCG 2579  
QY 1015 ggtgagtaactgactctctctgctatcaagtcacacaaacacttaactggtgttc 1074  
Db 2580 GGTGACCTTCTGCT 2639  
QY 1075 ttaaaagctcgacatcgcctaaatcgtatggtccggacttcagatccctatgtaaa 1134  
Db 2640 TTAAAGCGCGGACCTTACGAAATGTGATGTGTGACTTTCACATCTCACTCAAA 2699  
QY 1135 gtgaacctgtacatcgtcaaaaagaagaatctccaagaagaagaactcgtgaaatgca 1194  
Db 2700 GTGAACCTGTACATGCGCAAGAGAGATCTTAAAGAGAGAGAGAGAGAGAGAGAGAG 2759  
QY 1195 accccaagacagtgatcaatgaagcttctgcttcgtatccttcgtatgagagctgaa 1254  
Db 2760 ACTCCCAACGCGAGTGTCAATGAATCTTGTCTTGTATTCCTTGTGACAGCTTGA 2819  
QY 1255 gatataagtgatgaatttctgatttgatctcgtgaagaggggtcccgaaatgaagtaac 1314  
Db 2820 GAAATCAGTGTGAATCTTGTGTTGATTTGTAAGGGGAGTCCGAATAGAGTGAT 2879  
QY 1315 gggcagctgactcttggtggtgcagcagaagaagcgtggtggaagaagcagaagaagctc 1374  
Db 2880 GGACGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2939  
QY 1375 tctgactaccccaagagaagaatgcaagtggaagtgctctctgtagatggttagaactc 1434  
Db 2940 TGTGACTTCCCGAGAGCAATTTGTAATGTGACATCTCTGTGATGTGTTAGCTTCT 2999  
QY 1435 agccttgatggaagcttaaaagtttactagaagaagaattcttcttctat 1494  
Db 3000 GGCATGAGCTGCACACTTAAAGATTTTACTAGCAAGAGCAATTTCTTCT--TAT 3056  
QY 1495 attgagttgaaagcttg--ggaatcaagcttactcttgc-----gtgct 1559  
Db 3057 ACAGCTGTGATGAGCTTGCAGCAGACAGAGCTACTTATTTATTTATTTATTTACTTAT 3116  
QY 1540 tgttctgctgaagaatggaatggaatggaatggaatggaatggaatggaatggaatgga 1598  
Db 3117 TGTGTTGTTGAGAGTGAATTAAGTAAACCAAAATAGTTGTAATGTGTAATGTAATTA 3176  
QY 1599 tgataattccctatattatagaagttgataaaatttcaataagataatcaat--c 1656  
Db 3177 TGAATATTTCTCTTTTATTAAGAGTGTGCGCAAGTTTCAATAGATTAATGATTTCC 3226  
QY 1657 tccctcaagattcaagatgataactagaatagatagacatttgaacatggtgca 1716  
Db 3237 TCTGTGTGTAATTAAGTGGGCAACCCAGCTGTATGAGCACCTTACCTTAACCTTGT 3296  
QY 1717 gaatcccaaatataaagtgacatctcaatctgaaacatgtaacaagaatgaatgcat 1776  
Db 3297 ACAGATGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3356  
QY 1777 taagatttaaaaagcaaaagatagatgctctgctggaatctcaatctatctcaag 1836  
Db 3357 TTAGAGTTAAATAGAGACACACCTTTTGTATGAATTAATCACTGTTCTTCAGG 3416  
QY 1837 tgggggaagaatcaatttcttcaatacaagaataactaaagaatgctctcaag----- 1890  
Db 3417 TTGGGGAGATAGGTTTAAATTAATCCCAAGATATTTCCCAATGCTCTCCACATCGGT 3476  
QY 1891 ----ttgatattaatctctcatgycaaatggtgtgtctgcatataaagatctc 1946  
Db 3477 TTTTCTTTTCTTAAATACGTGATGTAACAAATGATGTTCTTATATATCAAGATATGC 3536



QY	1947	gagcatttcaggttggtttggtaattatttggatgcgaatttatcaaga---gtaacta	2003
Db	3537	TGT-----AATTTCATGGTGTATTTATTTGATACAGTCTCATCAAGACGGTAGCTTA	3590
QY	2004	gattcatctcaaaaggacagtgacaagctggaatattttat--caaaaggctgagt	2061
Db	3591	GGTTCATTTCATATAATATACACAAACAGTTGAGAAATTAATTTATACAGGCTGGAAT	3650
QY	2062	tgagaactgtgctggaatabaattttctcccccctaagttacatgtgagtc meta	2121
Db	3651	GAAACACACGTGGTGCATGATGAGATTTTTTTTTTTGTCCAA---TCACATGAGATCA---	3704
QY	2122	tttggtaaatataaactcaactaaagaacatggcttgatatttcaactccctgtaca	2181
Db	3705	---TGTAATAATTTAACTTCACTCACTAAGAACAGTGGCTGGAATTTATTCACCTCCGTGAGA	3761
QY	2182	agccctcagttgagccctggaanaatcccatgtaaccttctgaaatgtgtgaattagttagt	2241
Db	3762	AGCCTCAGTGTGGCCTAAGATACCTCTATTTCACATTTGGAAGTTGTGCAGCTAGTGA-	3819
QY	2242	gaataaagaataaacttcaactcagaanaatcccgcttagaagtgcaatttctctaagaaa	2301
Db	3820	-----AAAAAATAAACCTTAAGTAGCAATTCAGTTAGGAGATACATTTTTTAAATATATAA	3875
QY	2302	taggtat-agtgtgcaagtgtactttttagagcatcgtttgcaccagaagtcggactggc	2360
Db	3876	AATGTATCACTGTACAAAGTGTATTTTCAGAGATTGTCTTCCGCCCAAGATGACACATGCC	3935
QY	2361	caactaagttcattcaattaat 2381	
Db	3936	TACTTAAGTTTTTAACCTACT 3956	

RESULT	10
RNU14398	
LOCUS	RNU14398 2060 bp mRNA ROD 21-JUN-1995
DEFINITION	<i>Rattus norvegicus</i> synaptotagmin IV homolog mRNA, complete cds
ACCESSION	U14398
VERSION	U14398.1 GI:550453
KEYWORDS	
SOURCE	Norway rat.

ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE AUTHORS	1 (bases 1 to 2060) Ullrich,B., Li,C., Zhang,J.Z., McMahon,H., Anderson,R.G., Geppert,M. and Sudhof,T.C. Functional properties of multiple synaptotagmins in brain
JOURNAL MEDLINE REFERENCE AUTHORS	Neuron 13 (6), 1281-1291 (1994) 95085772 2 (bases 1 to 2060) Sudhof,T.C. Direct Submission
TITLE	Submitted (06-SEP-1994) Thomas C. Sudhof, Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235-9050, USA
FEATURES	Location/Qualifiers

```

source      location/qualifiers
1. 2060
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone_1lb="Strategene lambda ZAP adult rat brain cDNA
library"
/tissue_type="brain"
/dev_stage="adult"
268. 1545
/note="Similar to protein encoded by mouse synaptotagmin
IV, GenBank Accession Number U10355"
/codon_start=1
/product="synaptotagmin IV homolog"
/protein_id="AA68519.1"
/db_xref="GI:550454"

```

BASE COUNT 559 a 474 c 480 g 547 t  
ORIGIN

Query Match	28.9%	Score 1124.6;	DB 10;	Length 2060;
Best Local Similarity	84.0%;	Pred. No. 1.2e-206;		
Matches 1306;	Conservative 0;	Mismatches 244;	Indels 5;	Gaps 3

QY	105	caggcagatcttcctcttaagaacctcggaacgaacgaacgaataaaatggtcccgcatc	164
Db	223	CTGGCTGGAGTTCCTTTCAGACCTCCGGACAGACACGACG-AAAATATGGCTTCATAC	281
QY	165	caccacgcgggaagaatttgatgaatcccccacagtggtggtgacatccatgcatcttg	224
Db	282	CACGACGCGCGTGGAAATTCGATGAATTCACAGAGTGGGGCATCTTCAGTCTTTGG	341
QY	225	ccctggtctcacagtcctctctctctgtgcatgcatctgtcagagaataatccaatgc	284
Db	342	CTGTGTCCTTCACATGTCGTCTCTTTGGCTGGATCTGCTGCACAGAAAGATCCGCTAAATC	401
QY	285	taacaaagctccctccatcacagaatttgcga tgcgtctaagggatgtgatatccctga	344
Db	402	CACACAGACTCTCTCATACAGATTTGTGACACGTGCTTTAAAGAGTTGATATCTACCCGA	461
QY	345	aaacctaaatagcaaaaaagaagtttggagcagatgataaaatgaaatgaataagcc	404
Db	462	AAACCTAGTGTAGCAAAAAAGAGTTTGGAGAGATGACAAAGTAGCAGCAAAACCGTAAAGC	521
QY	405	agctgtgcgaagaatctcatctgcagctcgtgatcttgaagaagagatctcaatggcaatt	464
Db	522	AGCTTTGGCTTCACCTTTTCCTCGCACCTGTGATCTCGAAGACCGAGACTTAATGGCAATTT	581
QY	465	tcocaaacccaacctcaaacctgcagctctctctgtatctgaaatgcgaaccgcgaagt	524
Db	582	CCCCAAACCAACCCCAAGGTGGCAGCTTTTGATGTGGAATGTGCACCCCAAGGT	641
QY	525	ctttttaagaagggaaaaaagatgcagttcccccgtgaggttaagtccagaactccct	584
Db	642	CTTTCCGGAGACGGAAAAAGGAGGCGCTCTCCCTCGAGAGCTTCAAGTCCACGACATTTCCCT	701
QY	585	tacttcagaagaagaagaagaagatgcgtgggaactctctctctctctgaataaact	644
Db	702	CACCTTCAGAGAGAAACAGAGGAAGCTGGGACACTCTCTCTTCTGTCTGAGATCAACTT	761
QY	645	cgagagaaagcatcttgatgcataatcaagaagaagccgtgctctgcccagcatgtatga	704
Db	762	CGAGAGAAAGCATTTGTGGTGAAACATCAAGGAAGCCAGGGCTTACAGCCATGATGA	821
QY	705	gcaatgcatagcctctgacccatatataccaatgacgaatccctccagagaagaagcata	764
Db	822	ACAATTCATGTGACCTTCAGCCATTCATCAAGATGAGCAATTTACCGAGAAAGACACCA	881
QY	765	agtaaaactagatgctgttgagaanaacctgga tccagctcttgaatgagagaccttaact	824
Db	882	AGTGAAGACCGAGATGCTGTAGGAAGAGCGCTGGACCCCGTGTTTGACGAACCTTCACATT	941
QY	825	ctatggaataacctacaccanaatccaagaatgtgctctgacatccacaatttgaatt	884
Db	942	CTATGGGCTCCCTTATCCCAACATCCCAAGAGCTGTCCCTGCACATTACCGTCTGAGTTT	1001
QY	885	tgcacagtttccaagagatgatatcatcttgggagaagtctcaattctctctctcggaattga	944
Db	1002	TGACAGGTTCTCAAGAGATGATGTCATGTGAGGAAGTCCGTTCCTTCCTTCAGGGATTTGA	1061
QY	945	atlaatcuaaagaaaatgtctaataagaagaagatcatcaagaagaatgttaagaatgc	1004

```
Db 1062 ATTGTCAGACGGGAAATGTTATATGACGAGAGATCATCAGAGAAATGCTAGAAAGTC 1121
Qy 1005 ttgaggaaggggaggaattgactgactctctctgcatcaatgacacacacacttaac 1064
Db 1122 TTCGAGACGGGGGAGTACTTGTCTCCCTCTGTTATCAGTCCCTCAACACGCTCAC 1181
Qy 1065 tctgtgtctttaaagctcgaacatctgcctaaactcgaatgtgtgtgcggacttcaatcc 1124
Db 1182 TGTGCTGTCTTAAAGCGCGGACCTACCGAAATCTGATGTGTGACTTTCAGATCC 1241
Qy 1125 ctatgtaagtaaaccttaccatgcccacaaagagaatctccaagaagaactcaatg 1184
Db 1242 CTACGTAAGATGATACCTGTACCATGCGCAGAGAGAAATCTCTAAGAAAGACACTACGT 1301
Qy 1185 gaagaatgcaccccaatgcagtgctcaatgcagctgtgtgtcttgaataatcccttga 1244
Db 1302 GAAAGAGTGCATCCACACGACGTGTCAAGCACTGTTGTCTTGTATATTCCTTGTGA 1361
Qy 1245 gggccttgaagataaagtgttgaatttctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1304
Db 1362 GAGCTTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1421
Qy 1305 tgaagtaatgcggagcttgaatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1364
Db 1422 TGAGGTATGCGGGGCTGTGCTGTGCGGCGCAGAGAGAAAGAGTGTGCGGCGCAGCTG 1481
Qy 1365 gaaagagatctgtgaactaccccaagagaacaaatlgcacaagtgtgcagctgtgtgtgt 1424
Db 1482 GAAAGAGATGCTGTGATCTCCACGAGACAAATGCTAAGTGGCATATGCTGTGTATGTCG 1541
Qy 1425 ttagcactcctgacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1484
Db 1542 TTACTCTTCCGGCCACACCTACCACTTAAGGTTTACTAGCGGAGAAAGCAATTTTCT 1601
Qy 1485 tctcttcatatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1543
Db 1602 TTCT--TAAACCCGATTTGATGCTGTGCGCCAGCAAGCTACTTTCATTTATTTGTT 1658
Qy 1544 gttgtagaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1603
Db 1659 TTGCTTGAAGTGTGATTAATTAATGAGACCAAGATAGCTGTAAATGTATATGATGA 1718
Qy 1604 attcccatattatagaagagtgatgaataatttcatatgaataatcaatcctc 1658
Db 1719 ATTTCCTCTTTTATATGAGCGCTGTGCGCAAGTTTCATTAAGATCTGTTCCTC 1773

RESULT 11
RATRSI 1576 bp mRNA ROD 24-MAY-1995
LOCUS Rattus norvegicus (clone PCR2) synaptotagmin IV mRNA, complete cds.
DEFINITION L38247.1 GI:598376
VERSION L38247.1
KEYWORDS immediate early gene; synaptotagmin; synaptotagmin IV
SOURCE Rattus norvegicus (strain New England Deaconess Hospital) (clone
library: lambda ZAP PC12+ library (L. Vician and I.K. Lim))
ORGANISM Rhesus homoeocyteoma cDNA to mRNA.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1576)
AUTHORS Vician,L., Lim,I.K., Ferguson,G., Tocco,G., Baudry,M. and
Herschman,H.R.
TITLE Synaptotagmin IV is an immediate early gene induced by
depolarization in PC12 cells and in brain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 92 (6), 2164-2168 (1995)
MEDLINE 95199312
FEATURES
source Location/Qualifiers
1..1576
/organism="Rattus norvegicus"
/strain="New England Deaconess Hospital"
```

```
/db_xref="taxon:10116"
/cell_line="PC12"
/clone="PCR2"
/orig="lib"=lambda ZAP PC12+ library (L. Vician and I.K.
Lim)
/usage_type="pneochromeytoma"
1..238
5'UTR
CDS
240..1517
/codon_start=1
/product="synaptotagmin IV"
/protein_id="AA67327.1"
/db_xref="GI:598377"
/translation="MAPITTSRVERDELPTVVGIFSAFGLVFTVSLPAMICQRRSAS
SNKTPYKRVHVLKGVDTYPERLSSKKRFGDDSEAKRKRAALPNLSLHLEKRLDN
LEYNFEKARVNVNKEAGLIPAMDQSMSTSDPYIKMTLDEKHKVTRVLRKTLDPV
FDEFTFVGVPYPHIGELSHFTVLSFDRSDRVIGEVLPVLSGIELSDGKMLMTRF
IKRNARKSSGREGELVSLCYOSTNTLTVVVLAARHLPRSDVSGLSDPYKVNLTVA
KKRISKKKTHVKCTPNFAVFNELFVFDIPCELSIEVERLVLDSEKRSNEVIGRLV
LGATAEQSGGCHWKKEICDFPRROIAKMHLCDG"
1514..1576
BASE COUNT 420 a 394 c 384 g 378 t
ORIGIN
Query Match 27.1%; Score 1054.2; DB 10; Length 1576;
Best Local Similarity 85.6%; Pred. No. 4.1e-193;
Matches 1184; Conservative 0; Mismatches 198; Indels 1; Gaps 1;
Qy 105 cagcagatttcccttcagcaccctcgcagacacacagcaataaattgctccgactac 164
Db 195 CTGCTGAGATCTTCTTACGACCTCGGACAGACACGACAG-AAACATGCGCTCCATCAC 253
Qy 165 caccagcgggagaagattgtgaataatcccaagctgtgtgtgtgtgtgtgtgtgtgtgtgt 224
Db 254 CACGACCGCGGATTCGATGCAATTCGCCACAGTGTGCGCATCTTCAGTGTGTGG 313
Qy 225 cctgtctctcagctctctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 284
Db 314 CTTGCTTCTACCTGCTGCTCTCTTCTGCTGTGATCTCTGCCAGAGATCGCTAAAGTC 373
Qy 285 taacaagactcctcacaagaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 344
Db 374 CAACAAGACTCTCCATACAAAGTTTGGACGTCGTAAAGAGTGTATCTACCCAGA 433
Qy 345 aaacttaataagaagaagattgtgaagaatgaataaataagaagaataaagcc 404
Db 434 AAACCTAAGTACAAAAAGAAAGTTTGGAGAGATGCAAGATGACGAAAGCGTAAAGC 493
Qy 405 agctgtgcacaagaattcatgtgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 464
Db 494 ACCTTGGCTTAACCTTCCCTGACCTGTGATCTCGAAGACGACACTCATGCGCAATT 553
Qy 465 tcccaaaaacactcaaacctgtgcagctcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 524
Db 554 CCCCAAAACCAACCCCAAGAGCTTGTGTGATCTGGAATAATGACCCCAAAAGCT 613
Qy 525 cttcttagaagggaagaagatcagcttccctcagagagtttaagtcgaactccct 584
Db 614 CTTCCTGGAGAGCAAAAGAGGCGCTCTCCCTGAGAGCTTGAAGTGCACACTTCCCT 673
Qy 585 tactcaagaagaagaagaagaagctgtgaactctctctctctctctctttagaatacaact 644
Db 674 CACTTCAAGAGGAAAGAAAGAGAGCTGGGACACTCTCTTGTCTTGAAGTCAACTT 733
Qy 645 cgaagaagaagcatttgtgtcaatcaagaagaagccgtgtgtgtgtgtgtgtgtgtgtgtgt 704
Db 734 CGAGAGAAAGCAATTTGCTGTGAACATCAAGGAAGCCAGGCGCTTACCGACCATGATGA 793
Qy 705 gcaagtgaatgacccctgaccatatacaaatgaagatccctcccaagaagaagataa 764
Db 794 ACAATCATGACCTCTGACCCCATACATCAAGATGACAAATCTTACACAGAGAAAGACAA 853
```

Oy	765	agaaaacctgagctgctcgagaaaaaaccttgatccagcttttgaatgaagacattcatt	824		
Db	854	AGTGAATACCGAGAGTGTCTGAGGAAAGACCGCTGGACCCCGTGTTTACGAAACCTTCAATT	913		
Oy	825	ctatgggaatacccttaacacccaatactcaagaatttggcctctgacttccaatttga	884		
Db	914	CTATGGGGTCCCTTATATCCCAACATCCAAAGACCTGTCCCTGCACCTTCAACGGTCCGAGTTT	973		
Oy	885	tgaacggtttccaagagatgatcatcattggggaaattccattccctctcgggaattga	944		
Db	974	TGACAGGTTCTCAAGAGATGATGTCATGTGAGGAAGTCCCTGCTCTTTCAGGAGTTGA	1033		
Oy	945	attctbgaagaaaaatggttaatgaataaggaatatcaagaagaaatgttgaagatc	1004		
Db	1034	ATTGTCAAGACCGGAAAAATGTTAAATGACACGAGAGATATATCAAGAAATGCTAAGAGTCC	1093		
Oy	1005	tccaggaacgggtbgaagttaactgatactctctctgctatcagtcaccaccaaacctcaac	1064		
Db	1094	TTTCGGACGGGGTGAACTTCTGGTCTCCCTCTGTTATTCAGTCCACTACAAACAGCTCAC	1153		
Oy	1065	tgtggttgctttaaagctcgacatctgacctaaatctgaatgtgtcggacttcaagatcc	1124		
Db	1154	TGTGTTGTTCTTAAAGCGCGGACCTACCGAAATCTGATGTGCTCGACTTTCAGATCC	1213		
Oy	1125	ctatgtaaaatgtaacctgtatccatctgccaaaaagaagaatctccaagaagaagatcatgt	1184		
Db	1214	CTAGCTCAAAATGAAACCTGTACCATGCCAAGAAAGAAATCTTAAAGAAAGACTCAAGT	1273		
Oy	1185	gaagaatgaaccccccaatgacgttgaatgaagacttcttcttgatatctcttga	1244		
Db	1274	GAATAAGTGCACCTCCCAACGACAGTGTTCACGAACCTGTTTGTCTTTGATATTCCTTGTGA	1333		
Oy	1245	gggccttgaagaatataaagtgttgaatttltggttttgatctbgaagggggtcccgaaa	1304		
Db	1334	GAGCTTGAAGAAATTAAGTGTGAATTTTGTGATTTTGTGATTTCTGAAGGGGATCCCGAA	1393		
Oy	1305	tgaagtaatcgggcagatgactctctgggtgcacgacgaaagaactgtgtgaagacatg	1364		
Db	1394	TGAGGTGATCCGGGGGGTGTGTCTCGGGGGCCACAGAAAGAAAGTGTGTGGGGGGCACTG	1453		
Oy	1365	gaagaagatctgtgactaccccccaagagaagaacttgcgaagtggacgttgctctgtatgg	1424		
Db	1454	GAGGAGATCTGTGTACTTCCCCAGAGACAAATTTGCTAAGTGAATGCTCTGTGATGG	1513		
Oy	1425	ttaagatcctcagccgtgagtgtgacatcaaaagtttactcaggcaagagaaaaatttct	1484		
Db	1514	TTAGCTTCTCGGCCACCAACGATCAACAATTAAAGGTTTTTACTAGGCGAGAACCAATTTTCT	1573		
Oy	1485	ttc 1487			
Db	1574	TTC 1576			
RESULT 12					
FC004291					
LOCUS	BC004291	2950 bp	mRNA		
DEFINITION	Homo sapiens, similar to synaptotagmin 11, clone MGC:10881				
ACCESSION	BC004291	IMAGE:3621175, similar to complete cds.			
VERSION	BC004291.1	GI:13279139			
KEYWORDS	MGC.				
ORGANISM	human.				
AUTHORS	Homo sapiens				
TITLE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	Strausberg, R.				
REMARK	Direct Submission				
	Submitted (01-MAR-2001) National Institutes of Health, Mammalian				
	Gene Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2500,				
	USA				
	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				

COMMENT

Contact: MGC help desk  
Email: cgaabs+email.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Buterfield,  
Susana Chan, Readman Kiryuwnski, Reta Kutsche, Oliver Lee, Soo  
Letitia Hsiao, Martin Kirywnski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pavan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeeadi, Jacqueline  
Schein, Duane Smalls, Michael Smith, Lorraine Speed, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAL, Plate: 13 Row: J Column: 19  
This clone was selected for full length sequencing because it  
passed the following selection criteria: hexamer frequency ORF  
analysis.

## FEATURES

### Location/Qualifiers

### Source

CDS

```

/codon_start=1
/product="Similar to synaptotagmin 11"
/protein_id="AAH04291.1"
/db_xref="GI:13279140"
/translation="MAEITNIRPSFDYSPVYALIGASLYVCYSYVFWSCCHQQA
EKRRNPPYKFIHMLKGISIYPEETLSKKRIKVRKRDGDGRGRGGRRNLIVDAEAG
LLSDKDDPRGSSGSCIDQLPKMDYGEKRIISLPSTPEGSKTSPSSPEEDMGLSS
LTFEVDYNPRFKALVYVIOEAGHLVPMDDTQOGSDPYIKMTIIPDKRHRYKTRKRT
LTPEVDETFEYFGIPYSOLODLVHLFSLDFRSDRDYIGEVWPLAGVDPSTGKQL
LTIIRIKRNOKICISRGELQVLSLYOVAOMVYLVAKRLPMDITGLSGNGYVYV
NVYGRRIAKKKRHVHKYKCLNLIIFNSFIYDITPTDILPDISIEFLVIDDRTKNEV
VGRILCAHSYTAGAGAHMREVCESPPKPAKKMHSLEY"

```

Query Match 9.38; Score 363.6; DB 9; Length 2950;

Matches 745; Conservative 0; Mismatches 539; Indels 18; Gaps 4;

Oy	138	cagcagcaaaaaaatggcttcgcgatccatccccacacggcggaagaagtatggatgaatccccaac	197
Db	171	CTCTGATTTACGCATCATGGCTGTGAGATCATCCAAATATCCGACCTAGCTTTGATGTGTCACCGGT	230
Oy	198	agtggctggagatctctcagtgcatctggccttgctcttcacagctctct-----ctctgc	251
Db	231	GGTGGCCGGCCTCATCGGGGCGCTGTGCTGGTGTGTGTGTCGTGGACCGCTTTGGT	290
Oy	252	atggatctgcctgctcagagaaaaatatctccaagctcaacagactcctccatcaagtctgt	311
Db	291	CTGGGCATGCTGTCCACCGACGAGCAGAGACAGACAAAGACCAACCATTCACAAATTTAT	350
Oy	312	gcacttgctcctaaggagcttgatlatcttacccctgtaaaacctaataagcaaaagaagtctgg	371
Db	351	TTCACATGCTCAAAAGGCATCAGCATTTACCCAGAGACCCTCAGCAACGAAGAAATCAT	410
Oy	372	agcagatgataaaaaatgtaagtcaagaataaagccagctgtgtccaaagaattcaatgcatct	431
Db	411	CAAAATGGCGAGACCAAAATGTCCTGTGGAGAGCGAAAGTGGACCTAGAACTGTGTGGT	470



Db	303	CTGGTCAATGCTGCCACCAGCAGCAGCAAGAAAGCAAGCAAGAGGCCACCATTCAAAGTTTAT	362
QY	312	gcacgtgcttgaaggaggttgatatacttacccttgaacaactaaatagcaaaaagagtttgg	371
Db	363	TCACATGCTCTCAAAAGGATCAGCATATATCCACAGACGACCTCAGCAACAGAAAGAAATCAT	422
QY	372	agcagatgatataaataatgaaglaaagaataagcagctgtgtccaaagaattcaatgtcatct	431
Db	423	CAAAAGTCGGGAGAGCAAAAGATGCTCTGGGAGGAGAGGTGAGCTAGGAACCTGTGGT	482
QY	432	ggaatctgaaaagaagagatcctaibygcaattttcccaaaacccaactcaacttgcag	491
Db	483	GGAGGACGACGAGAGGCTGGCTGCTTAGCCGAGACAAGATCCACAGGGGCTTAGCTTGG	542
QY	492	tcctctgcatctygaataygaacccccgagctcttttgaagaagggaaaaagatcagt	551
Db	543	ATCTTGATATAGCAACATTATACCATTCAAATATGCACATATGGGGAAGAACTATAGGACCCAT	602
QY	552	ttcccttgaagatttaaatgltcag-----cacttccttacttcaagaagaacaaga	605
Db	603	TACAAAGCTTACCCCTGGGGAGAGCAAAACCAACCTTCATCATCTCCAGAGGAGATGT	662
QY	606	gaagctygggaactcttcttcttccctagaatacaacttgcagagaaaaagcatlttgt	665
Db	663	CATGCTAAGATCCCTCACCTTCATGTCATGTGACATTAATCTCCGAAAAAGCCCTGTGGT	722
QY	666	caaatataaagaagcccgtygcttgcagccatgabatgaagctgatgacttgcagacc	725
Db	723	GACATCTACAGGAGGCCACGCGCTGCCATATGATGGTGGACCAAGACCCAGGATCTGACCC	782
QY	726	atatataaaatgacgacatccctccagagagaagaataaagaagyaaaactaagatgtcag	785
Db	783	CTACATCAAAATGACCATCTCTTCTGACAAAGGATCGGGTGAAGACACAGTGTGCTCG	842
QY	786	aaaacacttgcataccagcttttgaatgaagaccttatactctatgtgataccctacaacca	845
Db	843	GAACACCTTGACCCCTGTGTTGATGACAGACCTTCACCTTCATGTCATCCCTCAGACCA	902
QY	846	aatccaagaatggccttgcagacttcaactcaatttggagtttgcaggttttcaagaagatga	905
Db	903	GCTTCAGGACCTGTGTGCTGCATCTCTTCTCTCAGACTTTGACCGCTTCTCTCGGAGTGA	962
QY	906	tatcaattggggaagtctcaattctctctcccyggaattgaatatctgaaggaaaaatgtt	965
Db	963	TGTCAATGGGAGGTCATGTGTGCCACTGTGCAGGGGGGAGCCACACAGGGGAAGGTCA	1022
QY	966	aatgataagaagatcaatcaagaagaatgttgaagatcttaaggacgggtgagttact	1025
Db	1023	ACTGACACGAGGACATCATCAAAAGGAATTCAGAAAGTGCATACACAGAGGGGAGCTCCA	1082
QY	1026	gactctctctgtatacaagcaccacaacacctcaacttgaatgtgtgtctttaaagctcg	1085
Db	1083	GGTGTCTCTCATATCAGCTGTGTGGCAGACGAAATGACAGTGGTGGTCTCTCAAAGCCAG	1142
QY	1086	acatctgcctaaatctgcatgtgtgcgagacttgcag---atcccatagtcgaagtgaacct	1142
Db	1143	ACACTTGGCGAAGATGGATATACACGGGTCTTCAGGTATCTTATGTGCAGAGTGAACGT	1202
QY	1143	gtacacatgcaaaaagaagaatctccaagaagaagactcatgtlgaagaatgcaacccaa	1202
Db	1203	CTACTACGGCAGAAAGCGCATTTGCCAAGAAAGAAAACCCATGTGAAGAAGTGCACATTGAA	1262
QY	1203	tgcagtgatgaagagctgtgttctttgatattctcttgtagggccttgaagataaag	1262
Db	1263	CCCCATCTTAATGAATCTTTCATCTACGACATCCCCCATGTGACCTCCGCTGATATCAG	1322
QY	1263	tgttgaatttggtttggatcttbaaagggggtcccgaaatgagatclogggcaatt	1322
Db	1323	CATGGAATTCCTCTTATTCGACTTTCGATCGCACACAAAGATGAGGTGTGGGAGGCT	1382
QY	1323	agctctgggtgcga---gcagcagaagaacctgtgtgagagcacttgaagaagatctgtga	1379

Db	1383	GATCCTGGGGGACACACATGTCACAGCCAGTGGTGCTCAACACTGGAAGAGGTCTGGCA	1442
Qy	1380	ctacccagagacaatttcgcaagtggcagctgtctgtga	1421
Db	1443	GAGCCCCGCAAGCCTGTGGCCAAAGTGCACAGTCTGAGCGCA	1484
RESULT	14		
LOCUS	AF000423		
DEFINITION	Rattus norvegicus synaptotagmin XI mRNA, complete cds.		
ACCESSION	AF000423		
VERSION	AF000423.1	GI:2130631	
KEYWORDS			
SOURCE	Norway rat.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 2426)		
AUTHORS	von Poser,C., Ichtchenko,K., Shao,X., Rizo,J. and Sudhof,T.C.		
TITLE	The evolutionary pressure to inactivate: A subclass of synaptotagmins with an amino acid substitution that abolishes calcium binding		
JOURNAL	J. Biol. Chem. (1997) In press		
REFERENCE	2 (bases 1 to 2426)		
AUTHORS	von Poser,C. and Sudhof,T.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-APR-1997) Molecular Genetics, UT Southwestern Medical Center, 5123 Harry Hines Blvd., Dallas, TX 75235, USA		
FEATURES	source		
	1..2426		
	/organism="Rattus norvegicus"		
	/db_xref="taxon:10116"		
	/tissue_type="brain"		
	243..1535		
	/note="membrane trafficking protein; syntaxin; intrinsc membrane protein with single transmembrane region and two C2-domains"		
	/codon_start=1		
	/product="synaptotagmin XI"		
	/protein_id="AAB58344.1"		
	/db_xref="GI:2130632"		
	/translation="MAETITNIRPSFDVSPVAAGLIGASVLVCSVTVFWTCCHOQA EKKKTPPYKRIHMLKGLSIYPERLISNKKELIKRPRQDSHRSGGNLLVNSGL LSHRDPRGPASPCINDLPSTIKRDYGELESPMSLPGSGSKPSPSPEDVMLGSL TFSVDYNPKKALVYITOENAGLPVMDGOTGOSPPTIKMTLPKRRTVTRVLKTL DPVDEFTTFPGIYSOLODLVLFHFLVLSFDRSRDVIQVWVPLAGVDSFGKVDL TRDILRNIOKCIISRGELQVLSYQVPAORNTVVVLAARHLPKMDITGLSGNPYKVN VYVYRKRILAKKTHVKKCTLNPIFNESFIYDIPDLPLDISIEFLVIDPRTTKNEVVA GRLLIGHSVTSSGAEMHREVCESPRKPVAKMHSLSSEY"		
BASE COUNT	616 a 610 c 628 g 572 t		
ORIGIN			
Query Match	8.9%; Score 346.2; DB 10; Length 2426;		
Best Local Similarity	56.9%; Pred. No. 8.4e-57;		
Matches	742; Conservative 0; Mismatches 538; Indels 23; Gaps 5;		
Qy	138	cacgcagtaaaaaatgtgtccgatcacaccacacgacggagaagaattgtatgaatcccaac	197
Db	230	CTCTGATTAAACAACATGCTGAGATCAACAATAATACGCCCACTTGTGATCTACACAGT	289
Qy	198	agtgttgaggatcttcaatgtcatctgtgacctgtgtcttcacagtctctc-----ctttgc	251
Db	290	AGCGGCTGCGCCATCATGGCGCCTGTGTCCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT	349
Qy	252	atgagtctgtctgtcagagaaatacatcacaagtctcaagaagactctcatatacaagtgtgt	311
Db	350	GTGACATGCTGCCACGACGACGCGGAAGAAGACACAAAGACCCACCTACAACTTAT	409
Qy	312	gcattgtcttaaggagattgatattacacctgaataatagcaaaagaatttgg	371

```

Db 410 TCATATGCTGAAGGACATTAGCATCTACCCAGACCCCTCAGCAACAAGAAATCAT 469
Oy 372 agcagatgaataaaatgaag-----taagaataagccagctgltccaaagatcat 424
Db 470 CAAAGTTCGGAGGACAAAGATGTTCTCATAGGCAAAAGTGGACGAGGAACCTGTTAGT 529
Oy 425 tgcactggaatcttgaagaagagatctcaatggaatcttcccaaaaccaactgaac 484
Db 530 AAATGCGGAGTCTGGCTCTGAGCATGACAGGATCCAGAGGGCTTACGCTCTCATC 589
Oy 485 ctggcagctctctgactcgcgaagaatgaaccccgaaactcttcttgaaggaggaaag 544
Db 590 TTGTATAGACCACTTACCCATCAAAAGAGACTATGGGAGAACTGAGAGAGCCCATGAC 649
Oy 545 agtgaagttccctggaagtttaagatccagactctcctactctgaagaagaagaag 604
Db 650 CAGCTTAACCCCGGGAAGCA----GCCACCTCCCATCATGCCAGAGAAAGATG 705
Oy 605 agaagctggaactctctctctctctctctctctctctctctctctctctctctct 664
Db 706 TGATCTAGGCTCCTTACTCTCTGCTGAGTACACTTCTCCAGAAAGCTCTGTGG 765
Oy 665 tcaatatacgaagaccgctgctgctgcaagcaatgagatgagatgactctgacc 724
Db 766 TGACATATCAAGAGGCGCATAGGCTGCGCATGATGATGAGGCCAGACCCAGGCTCCGATC 825
Oy 725 catatacaaatgacgactctctctctctctctctctctctctctctctctctctct 784
Db 826 CCTACATCAAAATGACCATTTCTCTGACAAACGCGCATGAGTAAAGACCCGGGTCTTC 885
Oy 785 gaaaaaacttgatccagcttcttgatgagacattacatctcagatgataccctacacc 844
Db 886 GCAAGACCCCTGAGCCCGGTGATGAGACCTTACCTCTACGCGATTCCTTACAGCC 945
Oy 845 aaatccaagaatgagctctgctctctctctctctctctctctctctctctctctct 904
Db 946 AGCTACAGGACCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1005
Oy 905 atatcatctggagagctctcaatctctctctctctctctctctctctctctctctct 964
Db 1006 ACGTACCTGGGAGGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1065
Oy 965 taagaataagaagatcacaagaagaagaatgtaagaagcttcaagcggggtgagttac 1024
Db 1066 AGCTGACCAAGGACATCATCAAGAGAACATTGACAGTGTAGCAGAGGGAATCTCC 1125
Oy 1025 tgaatctctctctctctctctctctctctctctctctctctctctctctctctct 1084
Db 1126 AGGATCTCTGTATACCGCCCTGCGCAGAGATGACAGTGTGTCTCTCAAGACCA 1185
Oy 1085 gaactctgctctaaatctgactgctgctgctgctgctgctgctgctgctgctgct 1141
Db 1186 GACATCTGCGGAAGATGATATACCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1245
Oy 1142 tgaatctctctctctctctctctctctctctctctctctctctctctctctctct 1201
Db 1246 TCTATCTAGGCAAAAGCAATTTGCAAGAAAGCAATGTAAGAAATGACATTTGA 1305
Oy 1202 atgcagctgctcaatgagctgctgctgctgctgctgctgctgctgctgctgctgct 1261
Db 1306 ACCCATCTTCAATGATGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1365
Oy 1262 gctgtaaatcttctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1321
Db 1366 GCATCTGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1425
Oy 1322 taactctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1378
Db 1426 TGATCTTGGGGGACACAGTGTACACACCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 1485
Oy 1379 acaaccacagagaagaatctgcaagctgcaagctgcaagctgcaagctgcaagctgca 1421
Db 1486 AGAGTCCCGCAAGCCTGTAGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1528

```

```

RESULT 15
LOCUS AB026808 1293 bp mRNA ROD 27-OCT-1999
DEFINITION Mus musculus mRNA for synaptotagmin XI, complete cds.
ACCESSION AB026808.1 GI:6136793
VERSION synaptotagmin XI.
KEYWORDS Mus musculus (strain:ICR) Adult Cerebellum cDNA to mRNA.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (sites) Kanno, F. and Mikoshiba, K.
AUTHORS Fukuda, M.
TITLE Conserved N-terminal cysteine motif is essential for homo- and heterodimer formation of synaptotagmins II, V, VI, and X
JOURNAL J. Biol. Chem. 274 (44), 31421-31427 (1999)
MEDLINE 2002669
REFERENCE 2 (bases 1 to 1293)
AUTHORS Fukuda, M.
TITLE Direct Submission
JOURNAL Submitted (27-APR-1999) to the DDBJ/EMBL/GenBank databases.
FEATURES
source source
location/Qualifiers
1..1293
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/dev_stage="Adult"
/tissue_type="Cerebellum"
1..1293
/codon_start=1
/product="synaptotagmin XI"
/protein_id="BA85780.1"
/db_xref="GI:6136794"
/translation="MAEINIRPSDVSVAAGLIGASVLVVCYSVTFVWTCCHQOA
EKKKRPVYFIMLKGISTVPELTSNKKIIVKRDGPRRSGNLLIYAESGL
LSHDXDVRGSPASCMDOIPIKRDGPELRSMPGSESPSPSPEDVWLSTL
TPSDVYRPPKALVYVTOBAHGLPYWDOTGSPYITMTLPPKRRVRYLRLT
DVPDDEPFYGTPIYSDLOLVLFHLYLSPDRSDYDYGWYPLMGWPSYKVL
TRDILRNKIQCSRGEIYVSIXQPVAGNNTVYLRKHLPKRDITGLSGNRYKVN
VYVRRRIARKKTHYKCTLVNPNVNSFTIDLPDLIDISTELVIDEPRTRNEV
GRLLDGHSVTTSGAHMEVCESPKPIAKMHSLEI"
BASE COUNT 337 a 347 c 346 g 263 t
ORIGIN
Query Match 8.9% Score 346; DB 10; Length 1293;
Best Local Similarity 57.1% Pred. No. 9,1e-57;
Matches 737; Conservative 0; Mismatches 530; Indels 23; Gaps 5;
Oy 151 atgagctccgataccaccagccggaagaattgcatgaaatcccaacagtggtggagtc 210
Db 1 atggctgagatccacaaatattacgccccagcttgatgctgacacggtagccgctgccc 60
Oy 211 ttcagtgcatctggcctgt-----cttcagagctctctctctctctgcatgcatctgt 264
Db 61 atccgggctctcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 120
Oy 265 cagagaaatcaccagaatcacaagaagccctccacataagaattgcatgcatgtctaa 324
Db 121 CACCAAGCAGGCGAGAGAAACCAAGACCCCGCATACAGTTCATTCATATGCTGAAA 180
Oy 325 ggaagtgtatttaccctggaagaacctaaatagaagaagaattggagcagatgataa 384
Db 181 ggcatttagcatcttaccagagaccctcagcaacaagaagaataatcgaagcttcgagca 240
Oy 385 atgtag-----taagaataagcagctgtgcaagaagaatcattgcatctgagatc 437

```

Db 241 GACAAAGATGGCCCTCGAGGGAAAGTAGCAGGACCTGTTATATAACGGGAGTCT 300  
QY 438 tgaagaagagatctcaatgycgaatttcccaaaacctcaacctgycagtcctc 497  
Db 301 GGCCTGCTGAGCCACGACCAAGGATCCAGAGGGCCCTAGCCCTGCACTCTGTATGAGACAG 360  
QY 498 tgaatcggaatgcaaccggaagctctttagaagggaagaaagagtcagttccccc 557  
Db 361 TTACCATTCAAAAAGAGACTATGGGGAAGACTGAGAGGCCCATGACAAAGCCCTGACCCCT 420  
QY 558 tgaagatttaaaagctcagcaactcccttaactcaagaagaagaaagaaagctggaac 617  
Db 421 GGTGAGAGCAAGCC---CACCTCCCATCTCGCCAGAGAAAGACGTATGTTAGCTC 476  
QY 618 tctctctctctccttaagaatacaacttgagaagaaagcatctgctgtaataataaga 677  
Db 477 TCTTACCTTCTCGTGACTACAACTTCCCAAGAAAGCTGTGTGTGACAAATCCAGA 536  
QY 678 agcccgctggtctgcccagcatgatgagcagtcgatgacctgacccatatacaaat 737  
Db 537 GGCCTCAGGGGCTGCCGATGATGATGATCAGACCCAGGGCTCCGATCCTACATCAAAAT 596  
QY 738 gaagatcctccagagaagaagataaagtgaanaactagatgctgagaanaaaccttga 797  
Db 597 GACCAATTCTCCGGACAAAGCGGATGAGAGACCCGAGTCTTCGCAAGACGCTGGA 656  
QY 798 tcaagctttagatgaacctttacattctatgagatacccttaacccaataccaagaatt 857  
Db 657 CCGGTGTTGAGTAGAGACTTCACTTCAAGGACATCCGTCACAGCCAGCTGAGGACCT 716  
QY 858 ggcctgcaactcaacaatttgagtttgaagagtttcaagagatgatacttgagg 917  
Db 717 GGTGCTGCAATTTCCTGTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 776  
QY 918 agttctaattcctctcctggaatgaattatctgaaggaagaaatgtaataagaga 977  
Db 777 GGTCAATGAGTCCATGCTGAGTGGACCCACACAGGCAAGGTCAGCTGACACAGGA 836  
QY 978 gatcatcaagagaatgatttagaagagctcgaagcgggtgaagttactgaatcctctc 1037  
Db 837 CATCATCAAGAGAGACATTCAAGAGTGCATTAAGAGAGGGAACCTCAGGTATCTCTGTC 896  
QY 1038 ctatcagtcacacacaacacttaactgltgltcttaaaagctcgacatctgctaa 1097  
Db 897 ATACAGCCCGTGCAGAGATGACAGTGGTGTCTCAAGCCAGACACTTGCCGAA 956  
QY 1098 atctgtagtgcggaacttcag---atccctatgcaagtgaaactgttaaccatgcaa 1154  
Db 957 GATGATATCACCGTCTCTCAGGTAATCCTTATGTCAAGGTGACGTCTATACGGCAG 1016  
QY 1155 aaagaagaatctcaagaagaagaagctatgtgaagaatgcaaccccaatgcaagtggtcaa 1214  
Db 1017 AAAACGATTGGCAAGAAAGAAACCATGTAAAGAGTGCATTGTGAACCCCGCTTTCAA 1076  
QY 1215 tgaagctgttcttctgataatctctgtagagagccttgaagataaagltgtaatttt 1274  
Db 1077 TGAAGTGTTCATATGACATCCCAACGACCTTGTGCTGATATCAGATCGAGTTCT 1136  
QY 1275 ggtttagatctgaaagggggtcccgaaatgaggtlaatcgggagttagctctggtgc 1334  
Db 1137 TGTCAATTGACTTCGATCGCACTACTAAGATGAGGTGTAGGAGGCTGATCTTGGGGGC 1196  
QY 1335 agcagcagaagagac---tggtagagaagccttgagaagagatctgtaactacccagag 1391  
Db 1197 ACACAGTGTACACAGAGTGTGCGAAGACCTGAGAGAGTCTGCGAAGAGTCCCGCAA 1256  
QY 1392 acaaatgccaatggcagctgctctgtga 1421  
Db 1257 GCCCATAGCCCAAGTGGACAGTCTGAGCGA 1286


Job time: 13057 sec

Fri Dec 14 10:31:38 2001

us-09-680-121-1.rge

---

Page 26





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2001, 10:53:59 ; Search time 3159.83 Seconds  
(without alignments)  
13232.302 Million cell updates/sec

Title: US-09-680-121-1

Perfect score: 3891  
Sequence: 1 cctctgctctcctcctgtt.....cccaaaaaaaaaaaaaa 3891

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: qb\_est1:\*  
11: qb\_est2:\*  
12: qb\_hic:\*  
13: qb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pin:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rnd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	718.6	18.5	1054	10	AV727109 AV727109
2	686.8	17.7	888	11	BG715690 602676962
3	646	16.6	656	10	AW160503 AW73612.Y
4	629.2	16.1	635	10	AW163709 au97d06.Y
5	625.2	15.9	723	11	BG699754 602681441
6	617.4	15.3	688	10	AV730176 AV730176
7	596.2	15.3	601	10	AW953983 EST365948
8	592.2	15.2	702	11	BG700203 602679821
9	564.8	14.5	591	10	AW162119 au73612.x
10	505.6	13.0	739	11	BG665092 DRABYC01
11	474.4	12.2	499	10	AI339352 q15f02.x
12	463.8	11.9	518	10	AA081755 zn22d07.r

13	458	11.8	565	10	AV606332 AV606332
14	445.2	11.4	712	10	AW914163 AW914163
15	443.4	11.4	475	10	AA484732 ne81f04.s
16	439.8	11.3	689	10	AW914162 EST345466
17	429.6	11.0	463	10	AA101663 AA101663
18	424	10.9	451	11	D53100 D53100
19	416.4	10.7	442	10	AI306404 AI306404
20	416.2	10.7	1174	11	BG261870 602373670
21	412.8	10.6	491	10	AA206761 AA206761
22	409.4	10.5	434	10	AA876133 cn33806.s
23	406.4	10.4	601	11	BF044050 BF044050
24	406.2	10.4	471	11	RS9082 RS9082
25	405	10.4	662	11	BF702882 MI-P-H2-a
26	402.2	10.3	494	10	AA206946 AA206946
27	399.2	10.3	449	11	H23524 H23524
28	391.6	10.1	398	10	AV728050 AV728050
29	387.4	10.0	435	10	AA084868 zn11f07.s
30	386.4	9.9	397	11	D52816 D52816
31	383.4	9.9	569	11	BF043853 BF043853
32	379	9.7	523	11	BG672430 BGC72430
33	378.8	9.7	525	11	H23552 H23552
34	376	9.7	379	11	Z44486 HSC21R031.n
35	369.4	9.5	491	11	RS1299 y972c07.s1
36	368.8	9.5	791	11	BG671553 DRNBUC05
37	368	9.5	369	11	Z43564 HSC1FB081.n
38	365.6	9.4	672	10	AV751605 AV751605
39	359.8	9.2	374	11	D56270 HUM422C08S
40	355.6	9.1	395	11	AA365639 EST76447
41	353.2	9.1	335	11	H05067 H05067
42	352.4	9.1	366	10	AA907386 OK95611.s
43	352.4	9.1	390	10	AA634047 ac34a05.s
44	351.4	9.0	376	10	AA977093 cq24a04.s
45	350.8	9.0	434	11	BF402440 UI-R-CAO-BF402440

#### ALIGNMENTS

RESULT 1  
LOCUS AV727109 1054 bp mRNA EST 17-OCT-2000  
DEFINITION AV727109 HTC Homo sapiens cDNA clone HTCAYE03 5', mRNA sequence.  
ACCESSION AV727109  
VERSION AV727109.1 GI:10836530  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1054)  
AUTHORS Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, D., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.  
TITLE Homo sapiens cDNA HTC clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@cnhg.sh.cn  
This clone is available at CHGC in Shanghai.  
FEATURES  
1. 1054  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HTCAYE03"  
/clone\_lib="HTC"  
/tissue\_type="Hypothalamus"  
/dev\_stage="Adult"  
/lab\_host="SOLR"

```

/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      322 a      189 c      200 g      330 t      13 others
ORIGIN

```

Query Match	18.5%;	Score 718.6;	DB 10;	Length 1054;
Best Local Similarity	86.2%;	Pred. No. 3.4e-131;		
Matches 904;	Conservative 0;	Mismatches 131;	Indels 14;	Gaps 10;

Oy	1153	aaaaagaagatctccaaagaagaagacatggaagaaatgacccccatgcagtgctc	1212
Db	1	AAAMAGAGATCTCCAAAGAAAGACATGTAAGAAATCCACCACATGACGTTC	60
Oy	1213	aatgaacctctgtctcttgatattctcttgaggcccttgaaagataagttgaaatt	1272
Db	61	AATGACCTCTGTCCTTTGATATTCTCTGTGAGGGCTTGAAGATATAGTGTGAATTT	120
Oy	1273	ctggtcttgatctcgaaggaggcccgaaatgagtaatcgggcagtgacttggtc	1332
Db	121	TTGCTTTTGATTTGGAAGAGNGGTCCCAAAATGAGTAATCGGGCAGTTAGCTTGCT	180
Oy	1333	gcgcagcagaagaagactcgttgagagcacctgaaagagatctgtactaccccgagga	1392
Db	181	GCAGCAGCAGAAAGAAACGTGTGAGAGCACTGGAAAGAGATCTGTGACTCCCCGAGGA	240
Oy	1393	caaatgccaagtgagcaagtgctcctgtatgtatgacatccaaagctgagtgtaacct	1452
Db	241	CAAAATGCCAAATGGCAGCTGCTCTGTGATGTTAGTACCTTACCCGTGAGTGAAGTT	300
Oy	1453	aaaggtcttactaagcaagaagaatacttctcttctatctgagtgcagctcgtg	1512
Db	301	AAAGTTTTTACTAGCAGAGAGAAATTTTCTTTCTTATTTGATTCGACGCTTG	360
Oy	1513	gaataaagctacccctttgt	1572
Db	361	GAATCAACACTACCTTTTGTGTCGTCTGCTGCTGTGAAATGATTAATTAAGTAC	420
Oy	1573	cagaagaagctcaaaagtgatcttgatatacttccctattctgaaagagttgata	1632
Db	421	CABAAGTACTTCAAAAGTGTATATGATAATTCCTATTTTATGAGAGCTTGATA	480
Oy	1633	aatttcaataagatatccaataatccctccagattaccagtgataaactaggaatgct	1692
Db	481	AACTTTCATTAAGATATTCATAATATCTCCCTGACATTAACAGAGATTAACTAAGAAATGCT	540
Oy	1693	agacattctagaactcgtgtgcaagaatcccaattaaatgagaaatctcatgtgaa	1752
Db	541	AGACATTTGATTAATAATACGTGCCAGATCCCAACATTAATAATGTAACAATCATATTGGA	600
Oy	1753	cagtgacaaaagaagttaaatgtatgaattt-aaaaagaaaagttgctt-gccttg	1810
Db	601	CAGTGACACAAAAAGTATGTGATTTACCATTTGATTAAGAAAGAAAGTATGCTTTGCTTGC	660
Oy	1811	tgaataattatcatcttactctcaagtttgagggaatacaattcttc--tltaaatcaag	1868
Db	661	TGAACATTAATCACTAATTAATCTTACGTTGGACAAATCACTTTCCTCAATTAATCAAG	720
Oy	1869	atcac-aaaaaaatgcctccagttgtgtatcttataatcttgcataatgtaagtgtgt	1927
Db	721	ATACCTAAAAAAATGTCGTCCAGTTTGATTTATTAATATCGTCATGTGAAT--GTTCC	779
Oy	1928	ccgtgcataaaagatcgtatcgtatcttggttgtgtaata--tttgaatcaattt	1986
Db	780	CCGCGATATACAA-TATCTGCTATTTCACTTTGTTGTTATTAATTAATCTTGTATCATCTTG	838
Oy	1987	atcataagatgaactca-gatcaattccaaaagaagtgacaagctgagaatattc	2045
Db	839	TTTATTAAGAAATGATCACCATTTATTTTATATGAGCCGTGAACACAGTTAANGACATATAT	898
Oy	2046	tatcaaaaggtctgagttgagaacacgtgtgcctgaaataa---tlttctccccccta	2101
Db	899	TTTATTAAGGCTTCAATTTGTAACTCTTGAGCATTAATAATTAATCTTTTNGCTCTCTTAA	958

Accession	Sequence	Position
Oy	aggtacacatgtagatgacaaatttctgtgaataataaacctgcacataagaacacatgcttgg	2161
Db	cgccacacatgtagacacaccttctgtgaataataaacctgcacatrtatcacctattggcc	1017
Oy	atatctacatgctgctgcacaaagcctaagt	2190
Db	attatctacatgtcttgcacaaacctcctaagt	1046

RESULT	2						
LOCUS	BCG715690						
DEFINITION	BCG715690	888 bp	mrna	EST	08-MAY-2001		
ACCESSION	6026765662P1	NIH_MGC_96	human sapiens	cDNA clone	IMAGE:4793930	5'	
VERSION	BCG715690						
KEYWORDS	BCG715690.1	GI:1399487					
SOURCE	EST.						
ORGANISM	human.						

CORRESPONDENCE  
NAME: Eukaryota, Metazoa: Chordata; Cranialata; Vertebrata; Euteleostom  
Mammalia; Eutheria: Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE: NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS: National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE: Unpublished (1999)  
JOURNAL: Contact: Robert Strausberg, Ph.D.

Email: [cgapds-r@mail.nih.gov](mailto:cgapds-r@mail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shira  
Toshliyski and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MOC Clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10688 row: 1 column: 15  
High quality sequence stop: 738.

```

FEATURES
source
1. 888
Location/Qualifiers
1. 888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4799390"
/clone_1lb="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript KS+; Site: 1: BamHI; Site 2: SalI; Site 3: XbaI; Site 4: PstI; Site 5: NotI; Site 6: KpnI; Site 7: SmaI; Site 8: XbaI; Site 9: PstI; Site 10: KpnI; Site 11: SmaI; Site 12: XbaI; Site 13: PstI; Site 14: KpnI; Site 15: SmaI; Site 16: XbaI; Site 17: PstI; Site 18: KpnI; Site 19: SmaI; Site 20: XbaI; Site 21: PstI; Site 22: KpnI; Site 23: SmaI; Site 24: XbaI; Site 25: PstI; Site 26: KpnI; Site 27: SmaI; Site 28: XbaI; Site 29: PstI; Site 30: KpnI; Site 31: SmaI; Site 32: XbaI; Site 33: PstI; Site 34: KpnI; Site 35: SmaI; Site 36: XbaI; Site 37: PstI; Site 38: KpnI; Site 39: SmaI; Site 40: XbaI; Site 41: PstI; Site 42: KpnI; Site 43: SmaI; Site 44: XbaI; Site 45: PstI; Site 46: KpnI; Site 47: SmaI; Site 48: XbaI; Site 49: PstI; Site 50: KpnI; Site 51: SmaI; Site 52: XbaI; Site 53: PstI; Site 54: KpnI; Site 55: SmaI; Site 56: XbaI; Site 57: PstI; Site 58: KpnI; Site 59: SmaI; Site 60: XbaI; Site 61: PstI; Site 62: KpnI; Site 63: SmaI; Site 64: XbaI; Site 65: PstI; Site 66: KpnI; Site 67: SmaI; Site 68: XbaI; Site 69: PstI; Site 70: KpnI; Site 71: SmaI; Site 72: XbaI; Site 73: PstI; Site 74: KpnI; Site 75: SmaI; Site 76: XbaI; Site 77: PstI; Site 78: KpnI; Site 79: SmaI; Site 80: XbaI; Site 81: PstI; Site 82: KpnI; Site 83: SmaI; Site 84: XbaI; Site 85: PstI; Site 86: KpnI; Site 87: SmaI; Site 88: XbaI; Site 89: PstI; Site 90: KpnI; Site 91: SmaI; Site 92: XbaI; Site 93: PstI; Site 94: KpnI; Site 95: SmaI; Site 96: XbaI; Site 97: PstI; Site 98: KpnI; Site 99: SmaI; Site 100: XbaI; Site 101: PstI; Site 102: KpnI; Site 103: SmaI; Site 104: XbaI; Site 105: PstI; Site 106: KpnI; Site 107: SmaI; Site 108: XbaI; Site 109: PstI; Site 110: KpnI; Site 111: SmaI; Site 112: XbaI; Site 113: PstI; Site 114: KpnI; Site 115: SmaI; Site 116: XbaI; Site 117: PstI; Site 118: KpnI; Site 119: SmaI; Site 120: XbaI; Site 121: PstI; Site 122: KpnI; Site 123: SmaI; Site 124: XbaI; Site 125: PstI; Site 126: KpnI; Site 127: SmaI; Site 128: XbaI; Site 129: PstI; Site 130: KpnI; Site 131: SmaI; Site 132: XbaI; Site 133: PstI; Site 134: KpnI; Site 135: SmaI; Site 136: XbaI; Site 137: PstI; Site 138: KpnI; Site 139: SmaI; Site 140: XbaI; Site 141: PstI; Site 142: KpnI; Site 143: SmaI; Site 144: XbaI; Site 145: PstI; Site 146: KpnI; Site 147: SmaI; Site 148: XbaI; Site 149: PstI; Site 150: KpnI; Site 151: SmaI; Site 152: XbaI; Site 153: PstI; Site 154: KpnI; Site 155: SmaI; Site 156: XbaI; Site 157: PstI; Site 158: KpnI; Site 159: SmaI; Site 160: XbaI; Site 161: PstI; Site 162: KpnI; Site 163: SmaI; Site 164: XbaI; Site 165: PstI; Site 166: KpnI; Site 167: SmaI; Site 168: XbaI; Site 169: PstI; Site 170: KpnI; Site 171: SmaI; Site 172: XbaI; Site 173: PstI; Site 174: KpnI; Site 175: SmaI; Site 176: XbaI; Site 177: PstI; Site 178: KpnI; Site 179: SmaI; Site 180: XbaI; Site 181: PstI; Site 182: KpnI; Site 183: SmaI; Site 184: XbaI; Site 185: PstI; Site 186: KpnI; Site 187: SmaI; Site 188: XbaI; Site 189: PstI; Site 190: KpnI; Site 191: SmaI; Site 192: XbaI; Site 193: PstI; Site 194: KpnI; Site 195: SmaI; Site 196: XbaI; Site 197: PstI; Site 198: KpnI; Site 199: SmaI; Site 200: XbaI; Site 201: PstI; Site 202: KpnI; Site 203: SmaI; Site 204: XbaI; Site 205: PstI; Site 206: KpnI; Site 207: SmaI; Site 208: XbaI; Site 209: PstI; Site 210: KpnI; Site 211: SmaI; Site 212: XbaI; Site 213: PstI; Site 214: KpnI; Site 215: SmaI; Site 216: XbaI; Site 217: PstI; Site 218: KpnI; Site 219: SmaI; Site 220: XbaI; Site 221: PstI; Site 222: KpnI; Site 223: SmaI; Site 224: XbaI; Site 225: PstI; Site 226: KpnI; Site 227: SmaI; Site 228: XbaI; Site 229: PstI; Site 230: KpnI; Site 231: SmaI; Site 232: XbaI; Site 233: PstI; Site 234: KpnI; Site 235: SmaI; Site 236: XbaI; Site 237: PstI; Site 238: KpnI; Site 239: SmaI; Site 240: XbaI; Site 241: PstI; Site 242: KpnI; Site 243: SmaI; Site 244: XbaI; Site 245: PstI; Site 246: KpnI; Site 247: SmaI; Site 248: XbaI; Site 249: PstI; Site 250: KpnI; Site 251: SmaI; Site 252: XbaI; Site 253: PstI; Site 254: KpnI; Site 255: SmaI; Site 256: XbaI; Site 257: PstI; Site 258: KpnI; Site 259: SmaI; Site 260: XbaI; Site 261: PstI; Site 262: KpnI; Site 263: SmaI; Site 264: XbaI; Site 265: PstI; Site 266: KpnI; Site 267: SmaI; Site 268: XbaI; Site 269: PstI; Site 270: KpnI; Site 271: SmaI; Site 272: XbaI; Site 273: PstI; Site 274: KpnI; Site 275: SmaI; Site 276: XbaI; Site 277: PstI; Site 278: KpnI; Site 279: SmaI; Site 280: XbaI; Site 281: PstI; Site 282: KpnI; Site 283: SmaI; Site 284: XbaI; Site 285: PstI; Site 286: KpnI; Site 287: SmaI; Site 288: XbaI; Site 289: PstI; Site 290: KpnI; Site 291: SmaI; Site 292: XbaI; Site 293: PstI; Site 294: KpnI; Site 295: SmaI; Site 296: XbaI; Site 297: PstI; Site 298: KpnI; Site 299: SmaI; Site 300: XbaI; Site 301: PstI; Site 302: KpnI; Site 303: SmaI; Site 304: XbaI; Site 305: PstI; Site 306: KpnI; Site 307: SmaI; Site 308: XbaI; Site 309: PstI; Site 310: KpnI; Site 311: SmaI; Site 312: XbaI; Site 313: PstI; Site 314: KpnI; Site 315: SmaI; Site 316: XbaI; Site 317: PstI; Site 318: KpnI; Site 319: SmaI; Site 320: XbaI; Site 321: PstI; Site 322: KpnI; Site 323: SmaI; Site 324: XbaI; Site 325: PstI; Site 326: KpnI; Site 327: SmaI; Site 328: XbaI; Site 329: PstI; Site 330: KpnI; Site 331: SmaI; Site 332: XbaI; Site 333: PstI; Site 334: KpnI; Site 335: SmaI; Site 336: XbaI; Site 337: PstI; Site 338: KpnI; Site 339: SmaI; Site 340: XbaI; Site 341: PstI; Site 342: KpnI; Site 343: SmaI; Site 344: XbaI; Site 345: PstI; Site 346: KpnI; Site 347: SmaI; Site 348: XbaI; Site 349: PstI; Site 350: KpnI; Site 351: SmaI; Site 352: XbaI; Site 353: PstI; Site 354: KpnI; Site 355: SmaI; Site 356: XbaI; Site 357: PstI; Site 358: KpnI; Site 359: SmaI; Site 360: XbaI; Site 361: PstI; Site 362: KpnI; Site 363: SmaI; Site 364: XbaI; Site 365: PstI; Site 366: KpnI; Site 367: SmaI; Site 368: XbaI; Site 369: PstI; Site 370: KpnI; Site 371: SmaI; Site 372: XbaI; Site 373: PstI; Site 374: KpnI; Site 375: SmaI; Site 376: XbaI; Site 377: PstI; Site 378: KpnI; Site 379: SmaI; Site 380: XbaI; Site 381: PstI; Site 382: KpnI; Site 383: SmaI; Site 384: XbaI; Site 385: PstI; Site 386: KpnI; Site 387: SmaI; Site 388: XbaI; Site 389: PstI; Site 390: KpnI; Site 391: SmaI; Site 392: XbaI; Site 393: PstI; Site 394: KpnI; Site 395: SmaI; Site 396: XbaI; Site 397: PstI; Site 398: KpnI; Site 399: SmaI; Site 400: XbaI; Site 401: PstI; Site 402: KpnI; Site 403: SmaI; Site 404: XbaI; Site 405: PstI; Site 406: KpnI; Site 407: SmaI; Site 408: XbaI; Site 409: PstI; Site 
```

```

Query Match      17.7%  Score 686.8  DB 11:  Length 888:
Best Local Similarity 93.4%  Pred. No. 6.2e-125:
Matches 796:  Conservative  0:  Mismatches 42:  Indels 14:  Gaps

Oy  2429  gatlttttcatctgcccagataaaagtgcttacttaaccacaacaacatgtaagctaca 2488
      |||||
Db    5  GATTTTTTACATGCCAGATAAAAGCTGTACTTACCCACCAACAAAGTGTAGACTACA 64
      |||||

Oy  2489  aatcgcttcagaagaacatctcaataattacatagcttcagaagaataatgcttaagc 2548
      |||||
Db    65  AAATGCTTCAGAGCAATCTTAATATTAATTTACATATGCTTCAGCAAAATATGCTTAAGC 124
      |||||

Oy  2549  tgcataattagacaacaaagaaatgcttctcactcttctttaggctaattgctctga 2608
      |||||

```

```

Db 125 TGTCAATTGACACAAAGAAATGTTTCATCTTTCTAGGCTAATTGTCCTGA 184
OY 2609 gctgtgtctatagaagcagtttaacagactgtgtctgtatcatcttccaagccaggg 2668
Db 185 GCTGTGTCTTATAGACACATGTTACAGACTGTGTCTGTATCATATTTCCAGTGCCAGG 244
OY 2669 tctgaatcattcaccgaacccgttgaatgaagccgaccccgctgataltaattgaaagaa 2728
Db 245 TCTGAATTCATTCAGAACCTGTTAGATTAAGCTGCACCCGCTGATTTATTTGAAAGAA 304
OY 2729 ttagcttgagagtaatgtcatatatttgagttccttagaagaatgagtggaacttgag 2788
Db 305 TTAGCTTGAGAGTAATGTCACTATATTTGAGTCTTAGAGAGATATGAGTGAACCTTGAG 364
OY 2789 tacacttaattatataaataatcagaagtagaataaagctcactgaaataattacattt 2848
Db 365 TACAGTGAATATTTAATTAATGCAAGTTAGAAATTAAGCTCACTGAAAATTTTACATTTT 424
OY 2849 gagtcaggttttgctgagtaactttagcagtttctgagaatgtgttgaatcacagtg 2908
Db 425 GAGTCAGGTTTGTGTGACGACTTATAGCAGTTTGGAGAAATGCTTTGATATCACAGTGT 484
OY 2909 ttgtaatctatgaagaaatgcatttccaaacactatatacagcttctttagactatg 2968
Db 485 TTGTAAATTCATGAAAATGCAATTTTCCAAACACTTATACATGCTTTTATGACTATG 544
OY 2969 cctaattgaagaaataatgatacttctgtatgtacaagaattaaataacacctctt 3028
Db 545 CCAATAGTAAGAAACATGATTTACATCTGTATGTACAAAGATTAATAATCAAC--TCTT 602
OY 3029 ttgtgctttaaataagacttggga-ttaaagaacataatcccaatcatt-gtcttca 3086
Db 603 TTGTGCTTTAAATGACATTTGGGATTTACAAAGCATATTTCCCAATCATTTGCTTTCA 662
OY 3087 ttccactac-aagatccactcagacatctgtctcactcggcactctcgt---gaaagc 3142
Db 663 TTCCACTTACAAAAGTCACCTCCACAGCATCTCTCCACTCCGCGCACTCTGTGAAAACG 722
OY 3143 aaccatgaatga-----ctgtagtagtggtgtagtttgaggaaagcaatgcccattt 3197
Db 723 AACATGAATATGACACTGTTTAAAGTGTTGTAACTTGGGCAAGTCAATATGCTTTT 782
OY 3198 atgatat-tgcatttgatatactgagccgctggaacagaaatataatgtgacacttcaaaa 3256
Db 783 AGCTCTGTTGCTTGGGCACTCAAGGCGCTTGACCAAGATATATTTTGGACCTCTGA 842
OY 3257 gttgtaaggggc 3268
Db 843 ATTGAGTAGGCGC 854

RESULT 3
LOCUS AW160503 656 bp mRNA EST 09-NOV-1999
DEFINITION au3c12.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone
IMAGE:2781910 5' similar to SM:SYT4_RAT P50232 SYNAPTOTAGMIN IV.;
mRNA sequence.
ACCESSION AW160503
VERSION AW160503.1 GI:6299536
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 656)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gaisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Stepce,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: au73c12.x1
Contact: Wilson RK

```

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40RP from Gibco  
High quality sequence stop: 427.

FEATURES  
source  
location/Qualifiers  
1..656  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2781910"  
/clone\_lib="Schneider fetal brain 00004"  
/sex="male"  
/tissue\_type="frontal lobe"  
/dev\_stage="5 months post-conception"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pBluescript SK (Stratagene);  
Site\_1: Sert; Site\_2: XhoI; Double-stranded cDNA was  
prepared from human fetal brain tissue. 5' and 3'  
adaptors were used in cloning as follows: 5' adaptor  
sequence:  
5'-GAGAGAGAGAGAGAGCTCAAGATCTTAATTAATTAATCCCCCCCCC-3'  
and 3' adaptor sequence:  
5'-GAGAGAGAGAGAGCTCAAGATCTTAATTAATTAATCCCCCCCCC-3'  
size-selected for >0.5 kb inserts and has an average  
insert size estimated at 1.2 kb. This library was  
constructed using the CAP-trapper method for full-length  
enrichment and has not undergone amplification. Library  
was constructed by Dr. Claudio Schneider (LNCIB-Area  
Science Park, Trieste, Italy). " 2 others

BASE COUNT 213 a 147 c 138 g 156 t

ORIGIN

Query Match 16.6%; Score 646; DB 10; Length 656;  
Best Local Similarity 98.9%; Pred. No. 6.8e-117;  
Matches 649; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

OY 137 aacagcagtaaaaaatggtccgatcacaccagccgggaagatttgtaaatccca 196
Db 1 ACACGCGATTAATAAATGCGTCCGATCACACAGCCGGGAAGATTGATGAATCCCA 60
OY 197 cagtggtgggacatctcagtgcaattggcctggtcttcacagtcctccttctgata 256
Db 61 CAGTGTGGGAGATCTTCAGTGCAATTGGCTGTGCTTCAACAGTCTCTTGGCATGGA 120
OY 257 tctgctgtaaggaataatcatcgaagcttaacaagaagctcctcatcaagttgtcatg 316
Db 121 TCTGCTGTGAGAGAAATATCATCCAAAGTCAAGACTCTCCATCAAGTNGTCCATG 180
OY 317 tgccttaaggagttgatatcttaccctgtaaaccttaataagcaagaagttggagcag 376
Db 181 TCGTTAAGGAGTGTGATATTATTTACCTGAAAACCTTAATACAAAAGAACTTTGGAGCAG 240
OY 377 atgataaaaaatgaagtaagaataagccagctgtgccaagaatctatgcatctgcatc 436
Db 241 ATGATTAATAATGAGTAAGTAAGTAAGCAAGCTGTGCTCAAGAAATTCATTCGATCGGATC 300
OY 437 ttgaaaaagagatctcatatgcaatttcccaaaaacactcaaacctgtgcagtcctt 496
Db 301 TTGAAAAGAGAGATTCATTCAGATGGCAATTTTCCAAAACCAACCTCAACCTGGAGCTTT 360
OY 497 ctgacttgagaatgaaccccgaaagctcttcttgaagagggaagaaagtgcaagttcc 556
Db 361 CTGATTTGAGATGACACCCCGAAGCTCTTTTGAAGAGGAAAAGAGTCAGTTTCCC 420
OY 557 ctgagaagtttaagtcagcaacttcccttactcagaagaagaacaagaagaagcttggaa 616
Db 421 CTGAGAGTTTAAGTCACACACTTCCCTTACTTCAAGAGGAAAAGAGAGAGCTGGAA 480

```

[illegible]

RESULT	4
LOCUS	AM163709
DEFINITION	Am163709 635 bp -mRNA EST 09-NOV-1999 au97d06.v1 Schaefer fetal brain 00004 Homo sapiens cDNA clone IMAGE:788203.5 similar to SW:SYT4_Rat P50232 SYNAPTOTAGMIN IV. ; mRNA sequence.
ACCESSION	AM163709
VERSION	AM163709.1 GI:6302742
KEYWORDS	EST.
SOURCE	human.

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Homio sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
111 (bases 1 to 635)					
Krimer, D., Kellen, M., Bowles, L., Dubnque, T., Geisla, C., Jost, S.,					
J., Kuhn, B., Schellenger, K., Steptis, M., Tan, F., Theising, B.,					
White, Y., Wylie, T., Waterson, R. and Wilson, R.					
Washu-MCI human EST Project					
Unpublished (1997)					
Other_ESTs: au97d06.x1					

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel. 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (image@lml.gov) for further information.  
Seq primer: -40RP from G1bco  
High quality sequence stop:434.

```
FEATURES      Location/Qualifiers
source        1. .635
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2784203"
/clone_1b="Schneider fetal brain 00004"
/sex="male"
/tissue.type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: Bluescript SK (Stratagene);
Site1: SstI; Site2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GACAGCAGACAGACGTCACGATCCTTAATTAAATTATCCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GACAGCAGACATCGATCTTTTTTTTTTTTTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LMCIB-Area
Science Par, Trieste, Italy)."
206 a 143 c 134 g 151 t 1 others

```

ORIGIN	Query Match
16.28;	Score 629.2; DB 10; Length 635;

Best Local Similarity 99.4%; Pred. No. 1.4e-113;  
Matches 631; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy	137	aacagccagtaaaaaatcgcctccgattccaccaccacccgaggaagaatttgcataatcccca	196
Db	1	ACACGCGAGTAAAAAATGGCTCCGATCCAGTACCACGACCGGGAAGATTTGATGAATATCCCA	60
Oy	197	caatcgatcgaggaatccttcagtgcatcttgagccgagtccttcacagctcctctcttcga	256
Db	61	CAGTGGGGGGAGTCTTCAGTGCATTTGGCTGGCTCTTCACAGTCTCTCTTTGCATGGA	120
Oy	257	tcgtcgtctcagagaaaaatcattccaaagctctaaagaaagctccctcctaaagtcttgcatg	316
Db	121	TCTCTGCTCGAGAAAAATCATCTCAAGTGTACAGAGCTCTCCATACAAAGTTTGTCATG	180
Oy	317	tgacttaagggaagtctgatttgccttcgaaaaacctaaataagaaaaagaatttgagcgag	376
Db	181	TGCTTAAGGAGAGTGTGATATNTACCTGAAAACTTAATAGCAAAAAGAAATTGGAGCAG	240
Oy	377	atgataaaaaatgaagttaaagaataagccaagctctgyccaagaattcatatgcatactgc	436
Db	241	ATGATTAATAATGATGAATAAAGATTAAGCCAGCTGTGCCAAAGAAATCTCTTGCACTGTGGATC	300
Oy	437	ttgaaaaagagaatctccaatggcaatttttcccataaacccaacttcaaacctcggagtcctt	496
Db	301	TTGAAAAAGAGATCTCAATGCAATTTTCCCAAAACCAACCTCAAACTCGCGCAGTCTT	360
Oy	497	ctgaatctcgaggaatgcaaaccccgaaagctcttttagaaggagaaaaagatgcagtttccc	556
Db	361	CTGATCTCGAGATGCAACCCCGAAGCTCTTTTGAAGGGGAAAAAGATCGTTTCCC	420
Oy	557	ctgaggaatttaaagtcagcaacttcccttaacttcgaagaagaacaagagaagctcgga	616
Db	421	CTGAGAGTTTAAAGTCCAGCACTTCCCTTACTTCAGAAAGAAACAAGAAAGCTGGGA	480
Oy	617	ctctctctctctctcttaagaatacaacttcagagaagaagaagcatcttggttcaataatgaag	676
Db	481	CTCTCTTCTTCTCTTAAGATACAACTTCGAGAGAAAGCAATTTGTGTATATATCAAG	540
Oy	677	aagccgcttgagcttcgacgcataatgataagcagtcgatgaacctctgacccatatacaaaa	736
Db	541	AAGCCCTGGCTGTCCAGCCCTTGATGAGCAGTGCATGACCTCTGACCCATATATCAAA	600
Oy	737	tgacgaatcctccagagaagaagaataaagtgaaa	771
Db	601	TGACGATCTCCACAGAGAAGAGCCCTTAAGTGAAA	635

RESULT	5
LOCUS	Bg699754
DEFINITION	Bg699754 723 bp mRNA EST 07-MAY-2001
ACCESSION	G6268144.1F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4814098 5', mRNA sequence.
VERSION	Bg699754
KEYWORDS	Bg699754.1 GI:13968377
SOURCE	EST.
ORGANISM	human. Homo sapiens Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 723) NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1998) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: MIXOS Palkovits, M.D., Ph.D. cDNA library Preparation: Michael J. Brownstein (NHGRI), Shiroki Toshiyuki and Piero Carninci (RIKEN) cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E Consortium/LLNL at:
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	



Qy	3100	tcaaccctcaagaacatcttctccatctggatctccgtgtaaagaacaatgtgaact	3159
Db	361	TCACCCACAGCATCTTCCTCCACTGGCATCTCTGTGAAGCAACATGAATGAACCTG	420
Qy	3160	agtaaggctgtatgatttgggaagatccaatgagccatttatgtatgacttggatcat	3219
Db	421	AGTAGGTGTGTAGTTGGGGAAGTCAAAATGGCAATTTATGTATGTGCATTTGGTATCAT	480
Qy	3220	gagggcgttggaacaagaatatatgttgagacctcgtaaagtgttaagggccaaactcaagt	3279
Db	481	GGGCGGTGGAACAGAAATATATGTGGACCTCTGAAAAGTTAAGTGGCCAAATTCTAAAT	540
Qy	3280	attctctcaaggagccggaagttaatggttggtagacgtctggaggtatgttgttggagcg	3339
Db	541	ATTCTTCACGGAGACCAAAATTATATGTGGACGACGTGATGTAATGTGTG-TGGACCGAG	599
Qy	3340	ggcgatatttttttttttttaacatlggaacaatlgaaacaacaacaataaattttaaattta	3399
Db	600	GCGGATTTTTCTTTTTTTTACATGTGAAACATTTG-AAACCAACAAACATTTGTAAATTA	658
Qy	3400	aatgatgatt 3406	
Db	659	AATGATGAT 665	

[illegible]

```
FEATURES
source
    Location/Qualifiers
    1. .601
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="MAGE_resequencences, MACC"
    /note="Vector: pBluescriptKsm"
BASE COUNT      200 a      103 c      107 g      191 t
ORIGIN
```

Query Match	15.3%	Score 596.2	DB 107	Length 601
Best Local Similarity	99.5%	Pred. No. 4,2e-10		
Matches 598	Conservative 0	Mismatches 3	Indels 0	Gaps 0
Oy	2031	gctgaggaattatcttcacaaagcgtggttgaaacacgtgctcggaataatttt	2090	
Db	1	gctgacacatttttttttttattcacaagcgctgattgaaacacgtgctcggaatttttt	60	
Oy	2091	tctcccccataaggtatcatgtgagtcacaaattttgttaaalatalaacctacataaagc	2150	
Db	61	ttctccccccttaaggttaccatgtatggttcaaaattttgttataatataaaccttaccataaagc	120	

OY	2151	catggcccttgattatctacatgcctctcaacagctcaagcttgagcttgaggcctggagaatcccat	2210
Db	121	catggcccttgattatctacatgcctctcaacagctcaagcttgagcttgaggcctggagaatcccat	180
OY	2211	gtaaccttggtaaaattgttgaattgcttgagataaagaataaactcaactagaat	2270
Db	181	gtaccttggtaaaattgttgaattgcttgagataaagaataaactcaactagaat	240
OY	2271	cgaattagaagctgcaattctctatagagaataggatagtgtagagtgacttttaag	2330
Db	241	cgaattagaagctgcaattctctatagagaataggatagtgtagagtgacttttaag	300
OY	2331	ggcaatgctttaccgacgaatgagacatggccaccgaatctcaataatattatgtcc	2390
Db	301	ggcaatgctttgtaccgacgaatgagacatggccaccgaatctcaataatattatgtcc	360
OY	2391	cgcgaagaagattagaatgacacacttgaagaagacttgaaagatttttaacatccagata	2450
Db	361	cgcgaagaagattagaatgacacacttgaagaagacttgaaagatttttttaacatccagata	420
OY	2451	aaagtgtaacttaacacaaacaaatgtaaacatcacaaaaatggttcaagagcaattcta	2510
Db	421	aaagtgtaacttaacacacaaacaaatgtaaacatcacaaaaatggttcaagagcaattcta	480
OY	2511	ataataattcaatatgttcaacgcgaataatgtcttagagctgttcaaatagacacacaaga	2570
Db	481	ataataatttactatgttcaacgcgaataatgtcttagagctgttcaaatagacacacaaga	540
OY	2571	atgtagtccaactacatcttcttaagagctaatgttgcttgagcgtgtgtctatagagcaatt	2630
Db	541	atgtagtccaactatcttcttcttaagagctaatgttgcttgagcgtgtgtctttaaagcagitt	600
OY	2631	a 2631	
Db	601	a 601	

RESULT	8
Bg700203	
LOCUS	Bg700203
DEFINITION	Bg700203 702 bp mRNA
	6026T9892.F1 NIT_MGC_95 Homo sapiens cDNA clone IMAGE:4812748 5'
ACCESSION	Bg700203
VERSION	Bg700203.1 GI:13969309
KEYWORDS	EST.
SOURCE	human.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominoidei: Homo. 1 (bases 1 to 702)				
NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				
Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgaps-r@mail.nih.gov">cgaps-r@mail.nih.gov</a>				
Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shira Toshiyuki and Piero Carninci (RIKEN)				
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc.				
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>				
Plate: L1AM10705 row: 1 column: 05 High quality sequence stop: 694.				

```

FEATURES
source
location/Qualifiers
1..702
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:4812748"
/tissue_type="hippocampus"
/lab_host="DH10B"

```

/note="Organ: brain; Vector: pbunescriptR (modified  
 Bluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgaag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to R0T 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 CAP-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIMH/NHRI, National  
 Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 222 a 111 c 138 g 231 t  
 ORIGIN

Query Match 15.2%; Score 592.2; DB 11; Length 702;  
 Best Local Similarity 97.1%; Pred. No. 2.5e-106;  
 Matches 677; Conservative 0; Mismatches 13; Indels 7; Gaps 7;

QY 2793 gtgaattataataatgaagtagaattagctactgaaaaattacatttgaat 2852  
 |||||||  
 Db 9 GTTGAATTTTAAATATGCAAGTTGAATTAAGTCTACTGAAAAATTTCATTTTGAGT 68  
 QY 2853 caggtttgtcgaactcttgagcagttttgagaatgtgttgatatacagtgctt 2912  
 |||||||  
 Db 69 CAGGTTTTGTGACAGTACTTACAGATTTTGAAGATGTGTGATATCACAGTGTGT 128  
 QY 2913 aaattcatgaaaaatgacattctccaaacattatagctcttttatgactatgccta 2972  
 |||||||  
 Db 129 AATTTCTATGAAAAATGCTTTTCCAAACACTTATACATGCTTTTATGACTATGCTTA 188  
 QY 2973 atgtaagaagaaatglatatcatctgtatgtacaagaatlaaaaaatcaactctttt 3032  
 |||||||  
 Db 189 ATGTAAGAAAAATGTATTCATTTCTGTATGTACAAAGATTAATAAATCAACCTCTTTTGG 248  
 QY 3033 tgcctttaaaatgactttggatataaaagcatattcccaatcatcttccatccac 3092  
 |||||||  
 Db 249 TCCCTTTAAATAGC-TTGGGATTTAAAAAGCATATTTCCCAATTCATCTTCATTCAC 307  
 QY 3093 tacaagtgactcaagacatctgtccactcgcagcatctgttgaagcaacatgaat 3152  
 |||||||  
 Db 308 TACAAAGTCACTCACAGCATCTTGCTCCACACGCGATCTGTGAAAGCAACATGAAAT 367  
 QY 3153 gaactgtagtagtgtgtagtttggggaagtcacatgaccatttlatgtatgtcat 3212  
 |||||||  
 Db 368 GAACGTAGTAGTGTTGTAG-TTGGGGAAGTCAATAGGCCATTTTATGATGTCATTTG 426  
 QY 3213 gtataatggccgttggaacagaatatagttggaccttgaaaagttgtaaggggccaa 3272  
 |||||||  
 Db 427 GTATCATGGGCGGTGAAACAGATATATGTGACCTTGAAAAAGTTGTAAAGGGCCAA 486  
 QY 3273 tctaaagtatcttcaagcagcagaagtaatgtgtgtgtagcagctggaagtatgt 3332  
 |||||||  
 Db 487 TCTAAGTATTTCTTCACGCGACGACCAAGATTAAATGTTGTAACGCTGAGGTATGG-TGTC 545  
 QY 3333 ggaagcagcagcagattttttttttaacatgga-aacaaatgaacacaaacaaatctt 3391  
 |||||||  
 Db 546 GGAAGCGCGCGATCTTTTAAAAACATGCAACAAATGAACCAACAAACATTTT 605  
 QY 3392 aaaaatt-aaaatgataattgttaaa-tagtttttagcttttaaaatttaagtgtttt 3449  
 |||||||  
 Db 606 AAAATTCACAATGATTAATTTGTAACGTGTTTAAAGCTTTTAAAAATTAAGTGTCTCC 665  
 QY 3450 gagtgtgaaaatgtgagt-aaactatgtgcaactgg 3485  
 |||||||  
 Db 666 GAGTGTGAACGTGTAGTCAACACTATTGTGACCTGG 702

RESULT 9  
 AW162119 591 bp mRNA EST 09-NOV-1999  
 LOCUS AW162119/c  
 DEFINITION au73c12 x1 Schneider fetal brain 00004 Homo sapiens cDNA clone  
 IMAGE:2781910 3', mRNA sequence.  
 ACCESSION AW162119  
 VERSION AW162119.1 GI:6301152

KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 591)  
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
 Krizman,D., Kuchta,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,  
 J., Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theising,B.,  
 White,Y., Wyllie,T., Waterston,R. and Wilson,R.

TITLE  
 JOURNAL  
 COMMENT  
 Other ESTs: au73c12..y1  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 458.

FEATURES  
 source  
 1..591  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2781910"  
 /clone\_11b="Schneider fetal brain 00004"  
 /sex="male"  
 /tissue\_type="frontal lobe"  
 /dev\_stage="5 months post-conception"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pbunescript SK (Stratagene);  
 Site.1: SstI; Site.2: XhoI; Double-stranded cDNA was  
 prepared from human fetal brain tissue. 5' and 3'  
 adaptors were used in cloning as follows: 5' adaptor  
 sequence:  
 5'-GAGAGAGAAAGAGCTCAAGATCTCTTAATTAATTAATCAACCCCCCCCCC-3'  
 and 3' adaptor sequence:  
 5'-GAGAGAGAGACTCGAGATTTTGTGTTTGTGTTT-3'. The library was  
 size-selected for >0.5 kb inserts and has an average  
 insert size estimated at 1.2 kb. This library was  
 constructed using the CAP-trapper method for full-length  
 enrichment and has not undergone amplification. Library  
 was constructed by Dr. Claudio Schneider (LNCiB-Area  
 Science Park, Trieste, Italy)."

BASE COUNT 199 a 118 c 75 g 199 t  
 ORIGIN

Query Match 14.5%; Score 564.8; DB 10; Length 591;  
 Best Local Similarity 99.3%; Pred. No. 6.1e-101;  
 Matches 588; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 3291 cagccagaagtaaatgtgtgtagcagctgagatgttgttggacgagccgatttt 3350  
 |||||||  
 Db 591 CAGCCGAAGTTATATATGTATGATGATGATGATGATGATGATGATGATGATGATGAT 552  
 QY 3351 tttttaacatggaacaatgaaacaaacaaacaaatctttaaaatttaaatgataatt 3410  
 |||||||  
 Db 531 TTTTAAACATGGAACAAATGAACCAACAAACAAATTTTAAATTAATTAATGATTAAT 472  
 QY 3411 tgttaaatgttttagcttttaaaatttaaatgttttggagtggtgaaaagtgtgataa 3470  
 |||||||  
 Db 471 TGTAAATAGTCTTTTACCTTTTAAATTTAAAGTGTGAGTGTGAGTGTGAGTGTAA 412  
 QY 3471 acctattgcaactgtgttcagaaagagaaagaaacaaacaaaggaattgaaacgagca 3530  
 |||||||  
 Db 411 ACTATTTCGACTGTTTTCAGAAAAAGAGAAAGAAACAAACAAAGGAATTTGAAACGGCA 352  
 QY 3531 gggagatcttaacataattcatcattctgcgaaaatgtaactgttttagaatgtatata 3590  
 |||||||

Db 351 GCGAGATCTTAATACCTAATTCATCATTTCTG-AAATGTAAGTATTGAAGATGATTA 293  
 QY 3591 caatataatgtaataatcttgaaatccgtgttacaatccgtgacgtatataacatgtaa 3650  
 Db 292 CAATATCAATGTAATATCTTGAATCTGTTACAAATCCGACATGATTTAAACAATGTA 233  
 QY 3651 attaatgtttgtcgtatgatacccaatccaccaccaatggaggtatatacgtttga 3710  
 Db 232 ATTAATGTTTGTCTATTAAGCAATCTCACCCCAATGGAGGTAAATGATGTTTGA 173  
 QY 3711 agaac-gtgaactcgtlaattgatttcttgaatgttgaactcaatagaagtgtttg 3769  
 Db 172 AGAAGTGTGTAACCTCAGTAATGATTTCTGTGATGTTGAACCAATGAGATGTTTG 113  
 QY 3770 gaaagaaacatggtgtgtgagacaatgctgttcttcttgagccagctcgtatgattt 3829  
 Db 112 GAAGGAGCATGTGTGTGAGACAGTGTCTGTTTGTGCGCAGCTGTATGATGTTT 53  
 QY 3830 gtaagaccatgtttgtaagacataataatgctgcttctgccccaaaaa 3881  
 Db 52 GTAAGACCATGTTTGTAAAGCATGAATGAATGCTGCTTTGGCCCAAAAAA 1

RESULT 10  
 BG665092 739 bp mRNA EST 30-APR-2001  
 LOCUS DRABYC01 Rat DRG Library Rattus norvegicus cDNA clone DRABYC01 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG665092.1 GI:13887014  
 VERSION EST  
 KEYWORDS Norway rat  
 SOURCE Rattus norvegicus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus

REFERENCE 1 (bases 1 to 739)  
 Xhao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G.,  
 Guo,H.S., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z., and  
 Zhang,X.  
 TITLE Distanced gene expression profiles of rat dorsal root ganglion  
 JOURNAL Induced by peripheral nerve axotomy  
 COMMENT Unpublished (2001)  
 CONTACT: Zhang Xu  
 LABORATORY of Sensory System  
 INSTITUTE of Neuroscience  
 320 Yue Yang Road, Shanghai 200031, P.R.China  
 TEL: 86-21-64748700-121  
 FAX: 86-21-64713446  
 EMAIL: xu.zhang@ion.ac.cn

FEATURES  
 source  
 1. 739  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="DRABYC01"  
 /clone\_lib="Rat DRG Library"  
 /sex="male"  
 /tissue\_type="dorsal root ganglion"  
 /dev\_stage="adult"  
 POLYA-No.

BASE COUNT 210 a 196 c 168 g 165 t  
 ORIGIN

Query Match 13.08; Score 505.6; DB 11; Length 739;  
 Best Local Similarity 84.28; Pred. No. 2.5e-89;  
 Matches 593; Conservative 0; Mismatches 109; Indels 2; Gaps 2;

QY 160 atccaccaccacccggagaaatttgatgaatcccccacagtggtggatctcaagta 219  
 Db 1 ATCACACACACCCGCTCAATTCGAAATTCGAAATTCGCCACAGTGTGCTCATGAGTCT 60  
 QY 220 ttggccggtgtcctcaagtcctcctcttgatgatactgctgcacagaagaatca-tc 278  
 Db 61 TTTGGCTGTGCTTCACTGTGTCTCTTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 120  
 QY 279 caagctacaagaactccatcaagaattgtgtcgtgtcgttaaggagttgatalta 338  
 Db 121 TAAGTCCAAAGACTCTCCATACAAAGTTTGACAGTGTGCTTAAGAGATTGATCTTA 180  
 QY 339 cccctgaacctataatagcaaaagaatttgagcagaatgaataaagaagtaaga 398  
 Db 181 CCCAGAAACCTAAGTACGCAAAAAGAAAGTTTGAGAGAGATGCACAGATGAAGCAAGCC 240  
 QY 399 taagcagctgtgccaagaatcattgatactgatacttgataagaagaagaatcattg 458  
 Db 241 TTAAGCAGCTTTGCTTAACCTTTCCTGCACTGTGATCTCGAAGAGCAGACCTCATG 300  
 QY 459 caatttcccaaaacacactcaactcagctcctctgatacttgagagatgaacc 518  
 Db 301 CAATTTCCCAAAACCAACCCCAAGCTGGACCTTCTGTGTAATGTCACCC 360  
 QY 519 gaagctctttagaagaaggaaaaagatcagttccctcgtgaagtttaagttcagac 578  
 Db 361 AAAGCTCTTCCGAGAGAGAGAAAGAGGCGCTGCCCTGAGAGCTTGAAGTCCACAC 420  
 QY 579 ttccttacttcagaagaagaacaagaagctgtggaactcctctcttcccttagaata 638  
 Db 421 TTCCCTCACTTCAGAGAGAAACAGAGAGAGGAGGACCTCTCTTGTCTTGAAGTA 480  
 QY 639 caacttcagaagaagaatctgtgtcatalcaatacaagaagccgtggtcagccat 698  
 Db 481 CAACTTCAG 540  
 QY 699 ggaatgagcagtcagtcagtcctcagccatataatcaaatgacagatcccccagaaga 758  
 Db 541 GGATGACACATCCATGACCTCTGACCATACATCAAGATGAGATGATCCAGAGAA 600  
 QY 759 gataaagtgaaaactagatgtcgtgagaaaacttgatccagcttcttgtagagacct 818  
 Db 601 GCACACAGTGAACACACAGTGTGTGGAAGAGCGTGCACCCGTTTGTAGCAAACTT 660  
 QY 819 tacattctatggaatacccaacccaatccaagaattgacct 862  
 Db 661 CACATTCATGAGGCT-CCCTTATTCACATCCAGAGAGTGTCT 703

RESULT 11  
 A1339352 499 bp mRNA EST 13-FEB-1999  
 LOCUS q15f02.x1 NCI\_CGAP\_GC4 Homo sapiens cDNA clone IMAGE:1947675 3',  
 DEFINITION mRNA sequence.  
 ACCESSION A1339352  
 VERSION A1339352.1 GI:4076279  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 499)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-r@mail.nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael





Db 240 TTGTGATGTCCTTAAGGAGGATTTGATATTTTACCCGAAACCTTAATAGCAAAAGAGT 299  
Oy 368 ttgagacaga tga taa aat tga aat aag aat aag cag cgt gtc tcc a a a t t c a t c t g c 427  
Db 300 TTGGAGACAGATGATATAAATGAAGTAAGTAAGCCAGCTGTG-CAAAGATTTTCTTGC 358  
Oy 428 atctgagctcttga a a a g a g a g a t c c a a t g c c a a t t t c c c a a a c c a a c t t a a c t g 487  
Db 359 ATCTGGATCTTTGAAGAGAGATCTCAATGGCAATTTTCCCAAAACCAACNTCAACCTG 418  
Oy 488 gcaatctctctgagcttga a a t g c a a c c c g a a g c t c t t t t g a a g g g a a a a g a g t 547  
Db 419 GCACTTTCTTGTGATCTGGAATGCAACCCGAAAGCTTTTGTGAAGGGGAGAAAGAGT 478  
Oy 548 cagttccctctgagat t t a a a g t c c a g a c t t c c c t 584  
Db 479 CAGTTT-CCTTGAGAGTTT-AAAGTCAGACATTNCT 513

RESULT 13  
AV606332 565 bp mRNA EST 30-AUG-2000  
LOCUS AV606332 Bos taurus kidney fetus Bos taurus cDNA clone EIKI034B06  
DEFINITION 5', mRNA sequence.  
ACCESSION AV606332 GI:9736705  
VERSION AV606332.1  
KEYWORDS EST.  
SOURCE COW.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 565)  
AUTHORS Sugimoto, Y., Hirotsune, S., Takasuga, A., Itoh, R., Jitohzono, A. and Suzuki, H.  
TITLE bovine cDNA sequencing  
JOURNAL Unpublished (2000)  
COMMENT Contact: Yoshikazu Sugimoto  
Animal Genetics Division  
Shirakawa Institute of Animal Genetics  
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan  
Tel: 81-248-25-5641  
Fax: 81-248-25-5725  
Email: kazusugi@cocoa.ocn.ne.jp  
Single pass sequencing.  
This clone was obtained from a polyA-deleted cDNA library.  
FEATURES  
SOURCE Location/Qualifiers  
1..565  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="Bos taurus kidney fetus"  
/tissue\_type="kidney"  
/dev\_stage="fetus"  
/lab\_host="DH10B"  
/note="Vector: pZRI; Site\_1: SalI; Site\_2: NotI; Poly A was deleted from a NotI site"

BASE COUNT 162 a 110 c 144 g 148 t 1 others  
ORIGIN

Query Match 11.8%; Score 458; DB 10; Length 565;  
Best Local Similarity 90.3%; Fred. No. 5.9e-80;  
Matches 501; Conservative 0; Mismatches 51; Indels 3; Gaps 1;

Oy 915 ggaagttcaatctctctcgcggaatgaatctcgaagaaatgttaagaatg 974  
Db 11 GGAAGTCTCAATCTCTCTGCGAGGAATGAATTAATGAAGAAATGTAAGAACG 70  
Oy 975 agagatcatcaagaagaatgttaagaatcttcgagacggggtgagttactgactctct 1034  
Db 71 AGAGATTACCAAGAAATGTAGGAATTTTCAGGAGCGGGTGAGTATTATGCTCTCT 130

Oy 1035 ctgcatacagtcaccacacaaacttaactgtgtgtgtcttaaaagctcgacatcgc 1094  
Db 131 CIGCTACAGCTGCTCACCAAAATACTGTACTGTGTGTTTAAAGCTCCGCACTGCC 190  
Oy 1095 taaactatgtgtcgcgacttcaagctccctatgtcaaaagtgaacctgtccatgcga 1154  
Db 191 TAACTGTATGTGTGTGACTTCAATCCCTATGCAAGTGAACCTGTACATGCCAA 250  
Oy 1155 aaagaatctccaagaagaactcatgtgaagaatcaccccaatgcagttcga 1214  
Db 251 AAAGCATCTCTAAAAGAGACACATGTGAAGAAATGACGCCCATGCGATTCMA 310  
Oy 1215 tgaagctgttcttgaatctctgtgaagggccttgaagataaagtgtgaattt 1274  
Db 311 TGAATCTTTGTCTTGTGACTTCTGTGAGGCTTGAAGAGATTAAGTGTGAATTTCT 370  
Oy 1275 gatttgaatctgaagggggtcccgaaatgaagtaacggcgagttagttcgtgtgc 1334  
Db 371 GATTGTGATCTGAAAGGAGATCCGAAATGAGTATGGCGGCTGTGCTGGAGC 430  
Oy 1335 agcagcagaagaactggtgagagcacttgaagaagatcttgaactaccacagagaca 1394  
Db 431 ACCACAGAGAGAGAGCTGTGAGAGCACTGGAAGAGATCTGTACTATCCAGAGCA 490  
Oy 1395 aattgcgaagtgcacgtctctgtgacttgaactcctacgcg---tgagttgaact 1451  
Db 491 AATCCCAAGTGGCGACGCACTGTGTGATTGCTTGTAGCACTTANCCGTGATGAATGAACT 550  
Oy 1452 taaagtttacta 1466  
Db 551 TAACGATTTTACTA 565

RESULT 14  
AW914163 712 bp mRNA EST 25-MAY-2000  
LOCUS AW914163  
DEFINITION RGIAB26 5' end, mRNA sequence.  
ACCESSION AW914163  
VERSION AW914163.1 GI:8079837  
KEYWORDS EST.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 712)  
AUTHORS Lee, N.H., Glodok, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
Kerlavage, A.R. and Adams, M.D.  
TITLE Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat  
Gene Index  
JOURNAL Unpublished (1998)  
COMMENT Other ESTs: EST345466  
Contact: Lee, NH  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3528  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
This clone is available through the ATCC, contact the ATCC  
tel#703-365-2700 for further information  
Seq primer: M13 Reverse.  
FEATURES  
SOURCE Location/Qualifiers  
1..712  
/organism="Rattus sp."  
/db\_xref="taxon:10118"  
/clone\_lib="RGIAB26"  
/clone\_lib="Normalized rat brain, Bento Soares"  
/note="Organ: Brain; Vector: pT73pac; Site\_1: EcoRI; Site\_2: NotI"

BASE COUNT 185 a 145 c 178 g 204 t  
ORIGIN




Fri Dec 14 10:31:54 2001

us-09-680-121-1.rst

---

Page 12



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2001, 10:54:04 ; Search time 320.99 Seconds  
(without alignments)  
10392.388 Million cell updates/sec

Title: US-09-680-121-1

Perfect score: 3891  
Sequence: 1 ctcttgcctcctcctcgtt.....ccccaaaaaaaaaaaaa 3891

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N.Geneseq\_1101.\*  
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
21: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT.\*  
23: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3891	100.0	3891	19	AAV54208
2	3891	100.0	3891	19	AAV57327
3	430.6	11.1	434	21	AAV57327
4	363.6	9.3	5310	22	AAV57327
5	362	9.3	2315	22	AAV57327
6	362	9.3	5517	22	AAV57327
7	361	9.3	424	20	AAV57327
8	264.2	6.8	318	16	AAV57327
9	216.6	5.6	383	22	AAV57327
10	216.6	5.6	383	22	AAV57327
11	216.6	5.6	383	22	AAV57327

C	12	177.4	4.6	936	22	AAV58252	Oligonucleotide D1
C	13	177.4	4.6	936	22	AAV58254	Oligonucleotide D1
C	14	177.4	4.6	936	22	AAV58257	Oligonucleotide D1
C	15	177.4	4.6	936	22	AAV58259	Oligonucleotide D2
C	16	177.4	4.6	936	22	AAV58262	Oligonucleotide D2
C	17	177.4	4.6	936	22	AAV58265	Oligonucleotide D2
C	18	173	4.4	936	22	AAV58252	Oligonucleotide D1
C	19	173	4.4	936	22	AAV58254	Oligonucleotide D1
C	20	173	4.4	936	22	AAV58257	Oligonucleotide D1
C	21	173	4.4	936	22	AAV58259	Oligonucleotide D2
C	22	173	4.4	936	22	AAV58262	Oligonucleotide D2
C	23	173	4.4	936	22	AAV58255	Oligonucleotide D1
C	24	163.2	4.2	1876	17	AAV29743	Mouse inositol pol
C	25	162	4.2	192	22	AAV19854	Probe #9787 for ge
C	26	162	4.2	192	22	AAV145050	Probe #13736 used
C	27	162	4.2	192	22	AAV10570	Probe #5561 used t
C	28	145.4	3.7	148	21	AAV25648	Human secreted pro
C	29	117	3.0	1569	22	AAV161195	Human polynucleoti
C	30	114.8	3.0	1647	22	AAV159409	Human polynucleoti
C	31	114.8	3.0	1660	21	AAV26182	Human ORF1737
C	32	112	2.9	4001	22	AAV57563	Human brain cell s
C	33	112	2.9	4001	22	AAV02800	Human shear stress
C	34	91.4	2.3	1337	21	AAV16073	Human prostate can
C	35	80.2	2.1	494	21	AAV09887	Human pancreatic c
C	36	80.2	2.1	528	21	AAV10589	Human secreted pro
C	37	79.2	2.0	4514	22	AAV18555	Human CDNA sequenc
C	38	71	1.8	72	21	AAV11525	Human secreted pro
C	39	70	1.8	2043	18	AAV79627	Human Doc2-beta ge
C	40	68	1.7	1718	17	AAV40760	Doc2 (brain-specif
C	41	68	1.7	1718	20	AAV72944	Human Doc2-alpha e
C	42	66.4	1.7	489	22	AAV05842	Human CDNA clone (
C	43	66.4	1.7	597	20	AAV88683	EST clone HC986.
C	44	64	1.6	244	22	AAV58238	Oligonucleotide D1
C	45	63.2	1.6	244	22	AAV58238	Oligonucleotide D1

## ALIGNMENTS

RESULT	1
ID	AAV54208 standard; CDNA: 3891 BP.
XX	XX
AC	AAV54208;
XX	XX
DT	11-JAN-1999 (first entry)
XX	XX
DE	Repro-PC-1.0 CDNA encoding prostate cancer-specific marker.
XX	XX
KW	Repro-PC-1.0; prostate cancer; marker: synaptotagmin; human;
KW	diagnosis; therapy: vaccine; ds.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	151..1428
FT	/*tag= a
XX	XX
PN	W09839447-A1.
XX	XX
PD	11-SEP-1998.
XX	XX
PF	06-MAR-1998; 98WO-US04488.
XX	XX
PR	15-MAY-1997; 97US-0047811.
PR	07-MAR-1997; 97US-0041246.
XX	XX
PA	(REPR-) REPROGEN INC.
XX	XX
PI	French CK, Schneider PA, Yamamoto KK;
XX	XX
DR	WPI: 1998-506363/43.
DR	P-PSDB; AAV74584.

XX New isolated Repro-PC-1.0 polynucleotide(s) - are used to develop  
 PT products for the detection and prophylactic and therapeutic  
 treatment of prostate cancer

XX Claim 4: Page 64-67; 87pp; English.

XX This cDNA sequence codes a prostate cancer-specific marker, termed  
 CC Repro-PC-1.0 (see AM74584), which represents a novel human brain  
 CC synaptotagmin isoform that may function in exocytosis and  
 CC endocytosis pathways. Subtractive hybridisation was used to  
 CC isolate Repro-PC-1.0 cDNA from a male LNCaP tumour cDNA library.  
 CC An overlapping clone (P55-1) was subsequently obtained from the  
 CC library using a probe containing 5' sequences of Rep-PC-1.0, and  
 CC the complete coding region was determined by sequencing P55-1 and  
 CC an overlapping RACE-PCR derived 5' and cDNA clone. The  
 CC Repro-PC-1.0 gene localises to chromosome 18. The invention  
 CC provides Repro-PC-1.0 polynucleotides, including probes and  
 CC primers, as well as a polynucleotide vaccine for eliciting an  
 CC immune response against Repro-PC-1.0. Also claimed are methods  
 CC for detecting Repro-PC-1.0 polynucleotides using the probes and  
 CC primers, a method of inhibiting Repro-PC-1.0 expression in a cell  
 CC using the antisense sequence, methods for diagnosing prostate  
 CC cancer and for detecting prostate cancer cells in a subject, for  
 CC following the progress of prostate cancer, for detecting a  
 CC chromosomal translocation of a Repro-PC-1.0 gene, and for  
 CC detecting polymorphic forms of Repro-PC-1.0.

XX Sequence 3891 BP; 1229 A; 671 C; 788 G; 1203 T; 0 other:

Query Match 100.0%; Score 3891; DB 19; Length 3891;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctcttgcctccctcctctgtccagagctggtgcctcgtgctgtgttttcag 60  
 Db 1 cctcttgcctccctcctctgtccagagctggtgcctcgtgctgtgttttcag 60  
 QY 61 cgttcgaaagccggtcgttgcagcaggaatggaatcgaagccaggtttccct 120  
 Db 61 cgttcgaaagccggtcgttgcagcaggaatggaatcgaagccaggtttccct 120  
 QY 121 cagcactcggacagacacagcagtaaaatggtctccgataccaccacggagaa 180  
 Db 121 cagcactcggacagacacagcagtaaaatggtctccgataccaccacggagaa 180  
 QY 181 ttgatgaatcccccaatggttggtggtctctcagtgcatcttgctgcttcacag 240  
 Db 181 ttgatgaatcccccaatggttggtggtctctcagtgcatcttgctgcttcacag 240  
 QY 241 tctctcttcagcagtcgtcgtcagagaataatcacaagtcatacaagatctctca 300  
 Db 241 tctctcttcagcagtcgtcgtcagagaataatcacaagtcatacaagatctctca 300  
 QY 301 tacaagtttgatgattggttlaaggagttgatatattacacctgaataaaccataagaca 360  
 Db 301 tacaagtttgatgattggttlaaggagttgatatattacacctgaataaaccataagaca 360  
 QY 361 tacaagtttgatgattggttlaaggagttgatatattacacctgaataaaccataagaca 360  
 Db 361 tacaagtttgatgattggttlaaggagttgatatattacacctgaataaaccataagaca 360  
 QY 421 tcatgtcatctgagatcttgaaaaagagatactcaatgagcaatttcccaaaccaactc 480  
 Db 421 tcatgtcatctgagatcttgaaaaagagatactcaatgagcaatttcccaaaccaactc 480  
 QY 481 aaacctggagctctctgatactggaatgcaaccccgagatccttttagaagggga 540  
 Db 481 aaacctggagctctctgatactggaatgcaaccccgagatccttttagaagggga 540  
 QY 541 aaagagtcagtttcccttgagagtttaagatccagcacttccttacttcagaagaga 600

Db 541 aaagagtcagtttcccttgagagtttaagatccagcacttccttacttcagaagaga 600  
 QY 601 caagagaagctgggaactctctctctcctccttagaatacacttcgagaagaaagcatt 660  
 Db 601 caagagaagctgggaactctctctctcctccttagaatacacttcgagaagaaagcatt 660  
 QY 661 gtgtcacaatcaaggaagcccggtggtgcagcagatgagatgagcagtcgaactct 720  
 Db 661 gtgtcacaatcaaggaagcccggtggtgcagcagatgagatgagcagtcgaactct 720  
 QY 721 gaccatataatcaaatgagacgactctccagagaagaagcatalaagtgaaaactagag 780  
 Db 721 gaccatataatcaaatgagacgactctccagagaagaagcatalaagtgaaaactagag 780  
 QY 781 ctgagaataaaccttggatccagcttltgagagaccttatactatgtggaataccctac 840  
 Db 781 ctgagaataaaccttggatccagcttltgagagaccttatactatgtggaataccctac 840  
 QY 841 acccaaatccaaagattggccttgcaacttcacaatttgagtttgacaagttttcaaga 900  
 Db 841 acccaaatccaaagattggccttgcaacttcacaatttgagtttgacaagttttcaaga 900  
 QY 901 gatgatatacttggggaagttctaatctctctcgtcggaattgaaatctcgaagga 960  
 Db 901 gatgatatacttggggaagttctaatctctctcgtcggaattgaaatctcgaagga 960  
 QY 961 atgttaatgaaatagagatcatcaatcaagaagaatgtaggaagttctcagagcgggtgag 1020  
 Db 961 atgttaatgaaatagagatcatcaatcaagaagaatgtaggaagttctcagagcgggtgag 1020  
 QY 1021 ttactgatactctctctcctacagtcacccacacacactcaactggtgtgttcttaaa 1080  
 Db 1021 ttactgatactctctctcctacagtcacccacacacactcaactggtgtgttcttaaa 1080  
 QY 1081 gctcgacactctgcttcaatctgagtggtccgagcttccagacccctcagataagtgaa 1140  
 Db 1081 gctcgacactctgcttcaatctgagtggtccgagcttccagacccctcagataagtgaa 1140  
 QY 1141 ctgtacacatgcgcaaaagagagatctccaaagaagaactcagtgtaaaagaaatgcacccc 1200  
 Db 1141 ctgtacacatgcgcaaaagagagatctccaaagaagaactcagtgtaaaagaaatgcacccc 1200  
 QY 1201 aatgcagtggtcacaatgagcgtgtgtcttgatattcctcttgagaggtccttgaaatata 1260  
 Db 1201 aatgcagtggtcacaatgagcgtgtgtcttgatattcctcttgagaggtccttgaaatata 1260  
 QY 1261 agtgtgaatcttctggtcttgatctcgtaaaggggtcccgaaatgaggttaatcggcag 1320  
 Db 1261 agtgtgaatcttctggtcttgatctcgtaaaggggtcccgaaatgaggttaatcggcag 1320  
 QY 1321 ttagtcttggttgagcagcagcaggaaggaactggtgagagcaatctgaaagaaatctgtgac 1380  
 Db 1321 ttagtcttggttgagcagcagcaggaaggaactggtgagagcaatctgaaagaaatctgtgac 1380  
 QY 1381 taccacaggaagacaatctgcgaagtgcagcagtgctgtgagtgatgacatccacagcgt 1440  
 Db 1381 taccacaggaagacaatctgcgaagtgcagcagtgctgtgagtgatgacatccacagcgt 1440  
 QY 1441 gagtggagacttaaggttttctatagcagaagagaatcttcttcttcttatatgga 1500  
 Db 1441 gagtggagacttaaggttttctatagcagaagagaatcttcttcttcttatatgga 1500  
 QY 1501 ttgcaagcttggaatacagcttcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560  
 Db 1501 ttgcaagcttggaatacagcttcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560  
 QY 1561 gaattagtagccagaagaacttcaaatgtgatatgataaattccattattatag 1620  
 Db 1561 gaattagtagccagaagaacttcaaatgtgatatgataaattccattattatag 1620  
 QY 1621 aagagttgataaaatttctcaatagatatcaatatctccttcgaattcccgatgataaa 1680

Db 1621 aagagtgtgataaatttccaagaatataatcccttcagattaccagtgatataa 1680  
QY 1681 cttaggaatagtcagacatttatgaatagtgtgccagaatcccaatttaattgtgaaa 1740  
Db 1681 cttaggaatagtcagacatttatgaatagtgtgccagaatcccaatttaattgtgaaa 1740  
QY 1741 atctcatcttggaacatgtccaaaagaattatgtatgaatttaaaaacgaagaatgtc 1800  
Db 1741 atctcatcttggaacatgtccaaaagaattatgtatgaatttaaaaacgaagaatgtc 1800  
QY 1801 ccttgctgttgaaaaattatccattatccttcaggttgagggaataatcatttccctta 1860  
Db 1801 ccttgctgttgaaaaattatccattatccttcaggttgagggaataatcatttccctta 1860  
QY 1861 atcccaagaattactaaaaaatgtccctcagtttgtatatttaattcctgtcatgtgcaa 1920  
Db 1861 atcccaagaattactaaaaaatgtccctcagtttgtatatttaattcctgtcatgtgcaa 1920  
QY 1921 tggttgcttcgcatataaagaatcctgtcatcttcagtttgtgttgtaattatgtatgc 1980  
Db 1921 tggttgcttcgcatataaagaatcctgtcatcttcagtttgtgttgtaattatgtatgc 1980  
QY 1981 aatttatcataagatgaactcagatltctcaaaaaggacagtgaacaagctgagaat 2040  
Db 1981 aatttatcataagatgaactcagatltctcaaaaaggacagtgaacaagctgagaat 2040  
QY 2041 tatttatcaaaaggtgtgtgttgagaacctgtgtgtgtaataatttttcccccct 2100  
Db 2041 tatttatcaaaaggtgtgtgttgagaacctgtgtgtgtaataatttttcccccct 2100  
QY 2101 aaggttacaatgtgagcaaaaattttgtaaaataaaactccaactagaataatccagttgaa 2160  
Db 2101 aaggttacaatgtgagcaaaaattttgtaaaataaaactccaactagaataatccagttgaa 2160  
QY 2161 gattatcaactcgtgtcccaagccctcagtggtgaggaatccctatgtacttgt 2220  
Db 2161 gattatcaactcgtgtcccaagccctcagtggtgaggaatccctatgtacttgt 2220  
QY 2221 gaaattgttgaattgtgtgtgaataaagaataaactccaactagaataatccagttgaa 2280  
Db 2221 gaaattgttgaattgtgtgtgaataaagaataaactccaactagaataatccagttgaa 2280  
QY 2281 gggcaatttcttataggaatagtgatgtgtgcaagtgtacttttaaggccatcgttt 2340  
Db 2281 gggcaatttcttataggaatagtgatgtgtgcaagtgtacttttaaggccatcgttt 2340  
QY 2341 gtaaccagaagtcggcagtcccaactaagtccttcaatttaattgtgccccagaaga 2400  
Db 2341 gtaaccagaagtcggcagtcccaactaagtccttcaatttaattgtgccccagaaga 2400  
QY 2401 ttaagaatgctacttgaagaagactgtgagaatttttaacattgcagataaagaatgtttac 2460  
Db 2401 ttaagaatgctacttgaagaagactgtgagaatttttaacattgcagataaagaatgtttac 2460  
QY 2461 ttaacaacaacaacaatgttaagaactcaaatcgttcaagaagacattcctaataattta 2520  
Db 2461 ttaacaacaacaacaatgttaagaactcaaatcgttcaagaagacattcctaataattta 2520  
QY 2521 catagttcaacgaaaaatatgtcttaagcgttcaaatltagacacaaaagaatgtttcca 2580  
Db 2521 catagttcaacgaaaaatatgtcttaagcgttcaaatltagacacaaaagaatgtttcca 2580  
QY 2581 ctatctttctagggcaatttgtctgaagcgttcttaagaagcgtttacagacttgt 2640  
Db 2581 ctatctttctagggcaatttgtctgaagcgttcttaagaagcgtttacagacttgt 2640  
QY 2641 gttctgtatcatcttcagtcagaggttctggaattcattcagaacccgttgatataa 2700  
Db 2641 gttctgtatcatcttcagtcagaggttctggaattcattcagaacccgttgatataa 2700  
QY 2701 ggtgcacccctgtgatatttgaaaaagaattagcttgagagtaatgtcacctataattgag 2760  
Db 2701 ggtgcacccctgtgatatttgaaaaagaattagcttgagagtaatgtcacctataattgag 2760

QY 2761 tcttagagaatgaatgaacttgatgaacagtgaatatttaaatatgcagaattgaa 2820  
Db 2761 tcttagagaatgaatgaatgaacttgatgaacagtgaatatttaaatatgcagaattgaa 2820  
QY 2821 atgaatctcatcttgaaaataattacatttgaatcaggtttgtgtccagtaacttagagtc 2880  
Db 2821 atgaatctcatcttgaaaataattacatttgaatcaggtttgtgtccagtaacttagagtc 2880  
QY 2881 tttagagaatgtgtttgatatacagtggttctgaatctcatatgaanaatgcatttccaaa 2940  
Db 2881 tttagagaatgtgtttgatatacagtggttctgaatctcatatgaanaatgcatttccaaa 2940  
QY 2941 caacttaicagtccttttatagctatgcctaaatgtgaagaanaatgatataattctgt 3000  
Db 2941 caacttaicagtccttttatagctatgcctaaatgtgaagaanaatgatataattctgt 3000  
QY 3001 tgtacaagaatlaaaaatacaacctcttttgtgtccttlaaaaatgacttttggaattaaa 3060  
Db 3001 tgtacaagaatlaaaaatacaacctcttttgtgtccttlaaaaatgacttttggaattaaa 3060  
QY 3061 agcatattcccaatcaatgtgttccatcccaagaagtcacccaagacatcctgtctc 3120  
Db 3061 agcatattcccaatcaatgtgttccatcccaagaagtcacccaagacatcctgtctc 3120  
QY 3121 cacttgccatctctgtgaaaagcacatgaatgaactgtatgtgtgtgtgtgtgtgtgtgt 3180  
Db 3121 cacttgccatctctgtgaaaagcacatgaatgaactgtatgtgtgtgtgtgtgtgtgtgt 3180  
QY 3181 agtcaaatgagccatttatatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtat 3240  
Db 3181 agtcaaatgagccatttatatgtatgtatgtatgtatgtatgtatgtatgtatgtatgtat 3240  
QY 3241 gttggaaccttgaaaagtgtglaagggccaatctaaatgtcccaagcagccagaag 3300  
Db 3241 gttggaaccttgaaaagtgtglaagggccaatctaaatgtcccaagcagccagaag 3300  
QY 3301 ttaatgtgtgtgagcaatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3360  
Db 3301 ttaatgtgtgtgagcaatgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3360  
QY 3361 tgggaacaatgaacaacaacaacaacattttaaatlaaaatgtacattgttaaatgt 3420  
Db 3361 tgggaacaatgaacaacaacaacaacattttaaatlaaaatgtacattgttaaatgt 3420  
QY 3421 ttttagccttlaaaattlaaagtgttttgaagtgtgaagaatgtgaataactattgtca 3480  
Db 3421 ttttagccttlaaaattlaaagtgttttgaagtgtgaagaatgtgaataactattgtca 3480  
QY 3481 actgtgtttcagaagaaggaagaagaacaagaagaattgaaacggcaggagatctt 3540  
Db 3481 actgtgtttcagaagaaggaagaagaacaagaagaattgaaacggcaggagatctt 3540  
QY 3541 aatacctaattcatcatcttctgcaaaagtgtactgttttgaagtatatacatatcaat 3600  
Db 3541 aatacctaattcatcatcttctgcaaaagtgtactgttttgaagtatatacatatcaat 3600  
QY 3601 gtgaatacttgaatccctgttacaacatcctgtcacatgtlatataacatgtlaaatgtgt 3660  
Db 3601 gtgaatacttgaatccctgttacaacatcctgtcacatgtlatataacatgtlaaatgtgt 3660  
QY 3661 tgtctgatatgccaatctccacccaagaatgggaggtatataactgttttgaagaagcgtta 3720  
Db 3661 tgtctgatatgccaatctccacccaagaatgggaggtatataactgttttgaagaagcgtta 3720  
QY 3721 actcggtaattgtattgtttgtatgtgttaactcaataagaaggttttgaagaagaagcat 3780  
Db 3721 actcggtaattgtattgtttgtatgtgttaactcaataagaaggttttgaagaagaagcat 3780  
QY 3781 ggtgtgtgagaacagtgtctgttcttctgtgtccagcctgtatgaatgttttgaagaccatgt 3840  
Db 3781 ggtgtgtgagaacagtgtctgttcttctgtgtccagcctgtatgaatgttttgaagaccatgt 3840

QY 3841 ttgttaagacataaataatctgtcttcccaaaaaaaaaaaaaaaaaa 3891  
 Db 3841 ttgttaagacataaataatctgtcttcccaaaaaaaaaaaaaaaaaa 3891  
 RESULT 2  
 ID AAV57327 standard; cDNA: 3891 BP.  
 AC AAV57327;  
 XX 21-DEC-1998 (first entry)  
 XX Hormone-regulated Repro-PC-1.0 gene.  
 DE Hormone-regulated Repro-PC-1.0 gene.  
 XX Repro-PC-1.0; prostate cancer; LNCaP; hormone-regulated gene;  
 KM human; synaptotagmin; ds.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 151..1428  
 FT /tag= a  
 XX MO9839661-A1.  
 XX 11-SEP-1998.  
 PD 06-MAR-1998; 98MO-US04519.  
 XX 15-MAY-1997; 97US-0047811.  
 PR 07-MAR-1997; 97US-0041246.  
 XX (REPR-) REPROGEN INC.  
 PA French CK, Yamamoto KK;  
 PI WPI: 1998-506379/43.  
 DR P-PSDB: AAW75782.  
 XX Identification of hormone-regulated traits, e.g. in cancers - by  
 PT exposing grafts of biological material to different hormonal  
 PT environments in animals of different reproductive states  
 PS Example 1; Page 59-62; 85pp; English.  
 XX This nucleotide sequence represents a hormone-regulated gene from  
 CC human prostate cancer cells. In order to isolate sequences that  
 CC are over-expressed in male LNCaP tumours, a male LNCaP-specific  
 CC probe was generated by 3 rounds of subtractive hybridisation with  
 CC female LNCaP tumour cDNA. The probe was used to perform a primary  
 CC screen of a lambda-ZAP male-LNCaP tumour cDNA library. Positive  
 CC plaques were subjected to secondary and then tertiary screens using  
 CC male- and female-specific probes to isolate partial clones  
 CC Repro-PC-1.0. Subsequent screening of the male-LNCaP tumour  
 CC library and RACE-PCR yielded a sequence containing a single 1275  
 CC open reading frame encoding 425 amino acids (see AAW75782). The  
 CC encoded protein has regions of homology to the C2 regulatory  
 CC domain of calcium-dependent isoforms of protein kinase C and to  
 CC isoforms of synaptotagmin. The gene was localised to chromosome  
 CC 18. The invention provides methods of identifying hormone-regulated  
 CC traits in a cell. The methods involve cultivating the cell as a  
 CC graft in 2 different hormonal environments and determining whether  
 CC expression of the trait differs in the 2 grafts. The methods can  
 CC be used to identify hormonally-regulated traits and hormonally-  
 CC regulated genes for use as targets for therapeutic intervention in  
 CC disease states, particularly cancers.  
 SO Sequence 3891 BP: 1229 A: 671 C: 788 G: 1203 T: 0 other:

Matches 3891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 cctcttgcctcctcctgttccagagctggtgcccctggctcgtcgtgtttcag 60  
 Db 1 cctcttgcctcctcctgttccagagctggtgcccctggctcgtcgtgtttcag 60  
 QY 61 cgttcggaagccggtcgttgaatccaggaagtgaatccagcagcagttccctt 120  
 Db 61 cgttcggaagccggtcgttgaatccaggaagtgaatccagcagcagttccctt 120  
 QY 121 cagcacctcgcagcaaacacgcagtaaaaaatggtccgcatcacccagccggaagaa 180  
 Db 121 cagcacctcgcagcaaacacgcagtaaaaaatggtccgcatcacccagccggaagaa 180  
 QY 181 ttgtatgaatcccccacagtggtggggtcttcagtgcatttggcctgtctcacagtc 240  
 Db 181 ttgtatgaatcccccacagtggtggggtcttcagtgcatttggcctgtctcacagtc 240  
 QY 241 tctctccttgatgatctgctgcagagaataatccaaagtctcaaaagaactcccca 300  
 Db 241 tctctccttgatgatctgctgcagagaataatccaaagtctcaaaagaactcccca 300  
 QY 301 tacaagtttgcatgtgcttaaggaggttgatattaccccgaaacctaaatagcaaa 360  
 Db 301 tacaagtttgcatgtgcttaaggaggttgatattaccccgaaacctaaatagcaaa 360  
 QY 361 aagaagtttgagcagatgataaaataatgaataagaataagcagctgtgcgaagaat 420  
 Db 361 aagaagtttgagcagatgataaaataatgaataagaataagcagctgtgcgaagaat 420  
 QY 421 tcaatggatcctggatccttgaaagaagatctcaatggaatttcccaaaacaaactc 480  
 Db 421 tcaatggatcctggatccttgaaagaagatctcaatggaatttcccaaaacaaactc 480  
 QY 481 aaacttgcaatcctccttgatctggaagaatgcacccggaagctcttttgaaggagaa 540  
 Db 481 aaacttgcaatcctccttgatctggaagaatgcacccggaagctcttttgaaggagaa 540  
 QY 541 aagaagtcagttccctcctggagagtttaagttcagcaattcccttaactgaagaagaa 600  
 Db 541 aagaagtcagttccctcctggagagtttaagttcagcaattcccttaactgaagaagaa 600  
 QY 601 caaggaagcctgggaactctctctctctcttaagatacaacttcgagaagaagcattt 660  
 Db 601 caaggaagcctgggaactctctctctctcttaagatacaacttcgagaagaagcattt 660  
 QY 661 gtgtcaatatacagaagagccgtgtgtccagcagatgagtcagtcagtcctct 720  
 Db 661 gtgtcaatatacagaagagccgtgtgtccagcagatgagtcagtcagtcctct 720  
 QY 721 gaccatatacacaataagagatcctcccgagaagaagacataaagtgaactagagtg 780  
 Db 721 gaccatatacacaataagagatcctcccgagaagaagacataaagtgaactagagtg 780  
 QY 781 ctagaagaaaacttgatccagcttcttgatgagacattacactctcagagataccctac 840  
 Db 781 ctagaagaaaacttgatccagcttcttgatgagacattacactctcagagataccctac 840  
 QY 841 acccaaatccagaagtgtgcttgcaattcacaatttgaatttgaagaagtttccaaga 900  
 Db 841 acccaaatccagaagtgtgcttgcaattcacaatttgaatttgaagaagtttccaaga 900  
 QY 901 gattgatcatctggggaagtctaatctcctcctcctgggaattgaatttcgaagagaaa 960  
 Db 901 gattgatcatctggggaagtctaatctcctcctcctcctcctcctgggaattgaatttcgaagagaaa 960  
 QY 961 atgttaatgaataagagatcatcaagaagaatgttagaagatcttagagcgggtgag 1020  
 Db 961 atgttaatgaataagagatcatcaagaagaatgttagaagatcttagagcgggtgag 1020  
 QY 1021 ttaactgatctctctcgtctatcagtcacccaacaacttaactgtgtgtcttaaa 1080  
 Db 1021 ttaactgatctctctcgtctatcagtcacccaacaacttaactgtgtgtcttaaa 1080

Query Match 100.0%; Score 3891; DB 19; Length 3891;  
 Best Local Similarity 100.0%; Pred. No. 0;



QY	2161	gattatctcactgctctgtcaaaagcctccagctgtggtcctgagaatccccaatgaccttgtc	2220
Db	2161	gattatctcactgctctgtcaaaagcctccagctgtggtcctgagaatccccaatgaccttgtc	2220
QY	2221	gaattctgttaattagttatgtgataaagaataaactcaactcagaatccagttisgaa	2280
Db	2221	gaattctgttaattagttatgtgataaagaataaactcaactcagaatccagttisgaa	2280
QY	2281	gtgcgaattctctataaggaataatagatagctgtgcaagtgtaactcttaagggcactgtt	2340
Db	2281	gtgcgaattctctataaggaataatagatagctgtgcaagtgtaactcttaagggcactgtt	2340
QY	2341	gtaccsagatcgagcatggtccacttaagctccatctaattatgttcccccagaanaa	2400
Db	2341	gtaccsagatcgagcatggtccacttaagctccatctaattatgttcccccagaanaa	2400
QY	2401	ttaagatgctactttgaaaagaagctgtaaagatttttaactgtgcagataaagaatgttac	2460
Db	2401	ttaagatgctactttgaaaagaagctgtaaagatttttaactgtgcagataaagaatgttac	2460
QY	2461	ttaaaccaacaacaatagttaagacttacaatatcggttccaagagcaattctaataattta	2520
Db	2461	ttaaaccaacaacaatagttaagacttacaatatcggttccaagagcaattctaataattta	2520
QY	2521	catatgtctacgcaaaataatagtcttagctgtcggtccaatctagcaacaagaagatgtgttca	2580
Db	2521	catatgtctacgcaaaataatagtcttagctgtcggtccaatctagcaacaagaagatgtgttca	2580
QY	2581	ctactcttctcagaagcctaattgtctgtgagctgtgtgtcgtatagagcagtttcaagactgt	2640
Db	2581	ctactcttctcagaagcctaattgtctgtgagctgtgtgtcgtatagagcagtttcaagactgt	2640
QY	2641	gtctctgtatcatcttcacgtgtccagaggtgtccgaattctcaagaacctgttaagttaaa	2700
Db	2641	gtctctgtatcatcttcacgtgtccagaggtgtccgaattctcaagaacctgttaagttaaa	2700
QY	2701	gtctgcacccctgtgatttttttggaaaagaattagcttgagagtgatgttcaactatattgagt	2760
Db	2701	gtctgcacccctgtgatttttttggaaaagaattagcttgagagtgatgttcaactatattgagt	2760
QY	2761	tctttagaagaatgaagatgtgaactctgagacagttgaattataataatgtcagaattga	2820
Db	2761	tctttagaagaatgaagatgtgaactctgagacagttgaattataataatgtcagaattga	2820
QY	2821	attaagttactgtgaaaaatttatacttctgagtcaggtttgtgtgcagtacttagcagtt	2880
Db	2821	attaagttactgtgaaaaatttatacttctgagtcaggtttgtgtgcagtacttagcagtt	2880
QY	2881	tttagagaatgtgtgtgtatcatcagagtggtttgttaaatctcattgaanaatgaccttttccaa	2940
Db	2881	tttagagaatgtgtgtgtatcatcagagtggtttgttaaatctcattgaanaatgaccttttccaa	2940
QY	2941	caactcttacaatgcttttttcttgactctgcctaaatgtgaagaanaatgttaactatctgt	3000
Db	2941	caactcttacaatgcttttttcttgactctgcctaaatgtgaagaanaatgttaactatctgt	3000
QY	3001	tgtacaagaattaaaaatcaacctctttttgtgctttaaataatgacctttggattaaaaa	3060
Db	3001	tgtacaagaattaaaaatcaacctctttttgtgctttaaataatgacctttggattaaaaa	3060
QY	3061	agcatattcccaatcatcttcttcaatccactacagaatgcaccacacagcatctctgtgc	3120
Db	3061	agcatattcccaatcatcttcttcaatccactacagaatgcaccacacagcatctctgtgc	3120
QY	3121	cactcagcatctctgtgaaagacaatgtgaactgtgagtagtgtgtgattgtgggaa	3180
Db	3121	cactcagcatctctgtgaaagacaatgtgaactgtgagtagtgtgtgattgtgggaa	3180
QY	3181	agtcgaatggccattttatgtatgtgcatttggtaatcagggccgttggaacagaatatat	3240
Db	3181	agtcgaatggccattttatgtatgtgcatttggtaatcagggccgttggaacagaatatat	3240



KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia; ss.  
 XX Homo sapiens.  
 OS  
 PN WO200153312-A1.  
 XX  
 XX 26-JUL-2001.  
 PD  
 XX  
 XX 26-DEC-2000; 2000WO-US34263.  
 PF  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 PR  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR P-PSDB: AM39577.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 PS  
 PS Claim 1: SEQ ID NO 936; 10078bp: English.  
 XX  
 XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AM38642-AM42213) with nootropic,  
 CC immunosuppressant and cytosaric activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 XX Sequence 5310 BP; 1421 A; 1237 C; 1236 G; 1416 T; 0 other;  
 SQ

Query Match 9.3%; Score 363.6; DB 22; Length 5310;  
 Best Local Similarity 57.2%; Pred. No. 2.8e-72;  
 Matches 745; Conservative 0; Mismatches 539; Indels 18; Gaps 4;

DB 478 tcaatgctccaaagagatcatgatataccagagaccctccagaaacaagaanaatcat 537  
 QY 372 agcagatgataaataatgaatgaagaataagccaagctgtgcgaagaattcatgtcatc 431  
 DB 538 caaagtgcggagagacaaagaatgtctctggagggaggtgtagcgtgaagaacctgtgt 597  
 QY 432 ggaatctgaaagagagatctcaatgtgcaatttcccaaaccaactcaaacctgtgcag 491  
 DB 598 ggaagcagcagagagctgtgcctgtctaaagccagaacaagatcccaagggtcagctctgt 657  
 QY 492 tccctctgactggaagaatgcaaccccgaaagctcttttagaaggagggaagatgcat 551  
 DB 658 atctgtatagaccaatctacccatcaaatgtgactatggggaagactaagagacctat 717  
 QY 552 ttcccttgagagtttaagtccag-----caatcccttaactcagaagaagaacaaga 605  
 DB 718 taacaagcttgaccctctgtggagagcaaacacacctctccatcatctccagagagagatgt 777  
 QY 606 gaagctgggaactctctctctctcttagaatacaactcagagagagaagcatctgtgt 665  
 DB 778 catgctagagatccctcaactctctcagtgagataactctccgaaaaaagccctgtgtgt 837  
 QY 666 caatataaggaagcccggtgtgtgtccagccatgagtgagcaagtcgatacctctgacc 725  
 DB 838 gacaatccagagagcccaagcggtgtcagtgatgagtgacacagccagagatctgacc 897  
 QY 726 atatatcaaaatgagcatctctcccaagaagaagatcaagtgaaactagatgtgtgtg 785  
 DB 898 ctacatcaaaatgacatccctctcttgacaacagcgcatcggtgtgagacagagtgctgtcg 957  
 QY 786 aaaaacttgatccagcttttgatagacctttatcatctctatagggatcccaacca 845  
 DB 958 gaagaccctggaccctgtgtgttgacagagaccttcaactctatagatccccaagcca 1017  
 QY 846 aatccaagaattgacctgtgacttcaacaattttgagttttagacaggttttcaagaagatga 905  
 DB 1018 gctgcagagcctgtgtgtgacttctgtccagcttgcagcttgcagcttctctcgtgagtga 1077  
 QY 906 tatcatgtgggaagcttcaattctctctctcgtggaattgaaattctcgaaggaanaatgtt 965  
 DB 1078 tgtcatgtgcgaaggtatgtgtgtgtcactgtgcaggggtgagcccaagcagaaggtataca 1137  
 QY 966 aatgaataagagatcatcaagaagaatgttagaagcttcagagcgggtgagtact 1025  
 DB 1138 actgaccaggaatcatcaataaagaatatccagaagtgatcaagaagggagagctcca 1197  
 QY 1026 gatctctctgtcatcagtgccacacaacactcaactaactgtgtgtcttaaaagctgtg 1085  
 DB 1198 ggtgtctctgtcatatcagctgtgtgcacagagatgacagatgtgtgtgtctcctcaagcag 1257  
 QY 1086 acatctgcttaaatctgagatgtgtcggactttcag---atccctatgtcaaatgtgaact 1142  
 DB 1258 acactgtgcgaagatgagatcatccagctgtcctcaagtaactctatgtcaagtgtgaagt 1317  
 QY 1143 gtaccatgccaagaagaagaatctccaagaagaagactcatgtgaagaatgtgaccccca 1202  
 DB 1318 ctactacggcagaagaagcattgtccaaagaagaacccaatgtgaagaatgtgacattgaa 1377  
 QY 1203 tgcagtgttcaatgagctgttcttcttgatatctctgttgagggtgcttgaagataag 1262  
 DB 1378 ccccatcttcaatgaatcttcaactcaatccacatcccaactgcagactcctcgtgatacag 1437  
 QY 1263 tgttgaattttgtgttttgatctctgaagaagggtgtccgaatgaatgtgagcaggt 1322  
 DB 1438 catcgagttcctctgtatcgaacttcgatccacacaagaatgaatgtgtgtgtgtgaggt 1497  
 QY 1323 agtctgggtgtga---gcagcagaagaagctgtgtgagagcactctggaagaagaagctgtga 1379  
 DB 1498 gatccggggggcacacagtgatcaagccagatgtgtgtgacactgtgagagagatgtctgtga 1557  
 QY 1380 ctacccagagagaacttccagatgtgcagtgctgtgtga 1421

Db 1558 gagcccccaagcctgtgagcaagtgcacagctctgagcga 1599

RESULT 5

AAH14256

ID AAH14256 standard; cDNA; 2315 BP.

AC AAH14256;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:11566.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INSTR.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX Claim 8; SEQ ID 11566; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences. AAB92446 to

XX AAB95893 represent human amino acid sequences, and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX Sequence 2315 BP; 585 A; 600 C; 570 G; 560 T; 0 other;

QY 138 caagcagtaaaatgctccgatacccaacccagccggagaagtattgataaattccccc 197

Db 183 ctctgattacgaacagcagcagatcccaataccgaactagcttgatgctcaccggt 242

QY 198 agtggaggagatcattgcatcttgccctggtctccacagctctctctcttcc 251

Db 243 ggtggcggcctcattcagcagcctctgctgctgctgctgctgctgctgctgct 302

QY 252 atgattctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 311

Db 303 ctgctcattgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 362

QY 312 gcatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 371

Db 363 tcaatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 422

QY 372 agcagatgataaaatgagtaaaatgagtaaaatgagtaaaatgagtaaaatgagta 431

Db 423 caaagtgcag 482

QY 432 ggtcttgag 491

Db 483 gtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 542

QY 492 tcccttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 551

Db 543 atcttgatagacaaattcccaataaagagagagagagagagagagagagagagagag 602

QY 552 ttccctgag 605

Db 603 tacaagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 662

QY 606 gaagctggag 665

Db 663 catctgag 722

QY 666 caatatacgaag 725

Db 723 gaaatctcag 782

QY 726 atatacaaatgag 785

Db 783 ctacatcaaatgag 842

QY 786 aaaaaccttgag 845

Db 843 gaagacccctgag 902

QY 846 aatcagaatgag 905

Db 903 gctcag 962

QY 906 taccattggag 965

Db 963 tgcattggag 1022

QY 966 atgaaatgag 1025

Db 1023 actgacacag 1082

QY 1026 gatctctcctcattcattcattcattcattcattcattcattcattcattcattcattc 1085

Db 1083 ggtgctcctcattcattcattcattcattcattcattcattcattcattcattcattc 1142

QY 1086 acatctgctcattcattcattcattcattcattcattcattcattcattcattcattc 1142

Db 1143 acacttgccgaagatgagatcaccggtctctcaggtacactcattgtaaggtgagcgt 1202

QY 1143 gtacacatgcaaaag 1202

Db 1203 ctactacgag 1262

QY 1203 tgcagtgatcag 1262

Db 1263 ccccatctcaatgatcatcttcatcagacatcccatcagccctccgcatatacag 1322  
 Qy 1263 tctgaatttctgttcttgatctcgaagggtcccgaaatggtatcgtggcagtt 1322  
 Db 1323 catcaggtccctgtatcagcttcgcaccccaagaatgagtggtggaggtc 1382  
 Qy 1323 agctctggtgca---gcagcagaagaactgtggaagacactggaagaatctgtga 1379  
 Db 1383 gatctctggggacacacaggtacacagccagtggtgctgaaactggaagagcttcgcga 1442  
 Qy 1380 ctaccacagagacaatgtccagtgacgctgtctgtga 1421  
 Db 1443 gaccccccgaagcctgtgccaagtgcacagctcgaagcga 1484  
 RESULT 6  
 AAI60519  
 ID AAI60519 standard; cDNA; 5517 BP.  
 XX  
 AC AAI60519;  
 XX  
 DT 22-OCT-2001 (first entry)  
 DE Human polynucleotide SEQ ID NO 4508.  
 XX  
 KM Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR MPI: 2001-442253/47.  
 DR P-PSDB: AAM41363.  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 PS Claim 1: SEQ ID NO 4508; 10078bp; English.  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemia and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 SQ Sequence 5517 BP; 1480 A; 1274 C; 1294 G; 1469 T; 0 other:  
 Query Match 9.3%; Score 362; DB 22; Length 5517;  
 Best Local Similarity 57.1%; Pred. No. 6,5e-72;  
 Matches 744; Conservative 0; Mismatches 540; Indels 18; Gaps 4;  
 Qy 138 cagcagtaaaaaatggctccgataccaccagccggaggaatgtgaaatccccc 197  
 Db 506 cctctgattacgacatggatgacacacataccgaactgaacttgatgtcacccgt 565  
 Qy 198 agtgtgggagatctcagtgacttggcctgtgttcacagactctctc-----cttgc 251  
 Db 566 ggtggccgcccctatcgggccctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 625  
 Qy 252 atggaatctgtctcagagaaaaatcccaagctcacaagaactcctccataaagtgt 311  
 Db 626 ctggtcatgtctccaccagcagcagagagaagacacagaacccacatacaagttaa 685  
 Qy 312 gcatgtgcttaaggaggttgatataatcccttgaaacccaataagaanaaagaatttg 371  
 Db 686 tcaactgttcaaaagatatacagataccacagaccctcagcaaaagaanaaataat 745  
 Qy 372 agcagatgtataaaatgaagtaaaagataaagcagctgtgccaagaattcatatgatct 431  
 Db 746 caaagtgcggagagaaagaagaatggtcttgggaggaagtgtgaagtaaacctgttgt 805  
 Qy 432 ggatcttgaagaagagatctcaatgtgcaatttcccaaaacccaacccaacctgtgcag 491  
 Db 806 ggaagcagcagagagctgctgtcgttaagccgagacaaagatcccaagggtgctgtgtg 865  
 Qy 492 tctctctgactcggagaatagcaaaccccgaaagctctttagaaggaggaagaagatcagt 551  
 Db 866 atctgtatagacaattacccatcaaatgtgactatgggaagtaactaaggagccctat 925  
 Qy 552 ttcccttgagagtttaaatccag-----cacttccttactcagaagagaacaaga 605  
 Db 926 tacaagcctgacccctggggagagagaacaacacacttccatcatcctccagagagatgt 985  
 Qy 606 gaagctgggaactctctctctcctttagaatacaacttcgagaagaagaacattgtgtg 665  
 Db 986 catgctagtagtccctcactctcactctcagtgagataacttcccgaaaaagccctgtgt 1045  
 Qy 666 caatatacgaagaagccgtgtgtgtccagccatgtatgtgacgtcgtaccc 725  
 Db 1046 gacaaactcggagggcccaaggtgtgacagtgatgtgacagcccaagagatctacc 1105  
 Qy 726 atatatcaaaatgagatctctcccaagaagaagcaataagttaaaactagatgtctgag 785  
 Db 1106 ctacatcaaaaatgacatctctccctcgaacaagcgcatcggtgtgaagacaagatgtctgag 1165  
 Qy 786 aaaaacttgatccagcttlttagatagaccttatacttcatgtgagataccaccacca 845  
 Db 1166 gaagaccctggagccctgtgtttagcagagacactcactctcattagatccaccacagcca 1225  
 Qy 846 aatccaagaattgctgtgaacttcaaaatttgagtttgacaggttttcaagagatga 905  
 Db 1226 gctcagagacctgt 965  
 Qy 906 tatcatgtggagagttcattatctctctccgggaatttgatattatcgaaggaanaattgt 965  
 Db 1286 tgcatttgcgaagttatgt 1345  
 Qy 966 aatgaaataagagatcatcaagaagaattttaggaagcttcagagcgggtgtgtgtgt 1025

Db 1346 actgaccaggacatcatcaaaaagatatccagaatgcatcagacagagggagctcca 1405  
 OY 1026 gatctctctctgtatcatgaccacaaacactctactgtgtgtcttaaaagctcg 1085  
 Db 1406 gggtctctctatcatcagcctgtgcccagagatgacagtggtgtgctccaaagccag 1465  
 OY 1086 acatctgctcaatctcgatgtgtcggacttcag---atccctatgtcaagtgaact 1142  
 Db 1466 acacttgcagaagaatgatatccgctccctccggaalcccttaagtcaagtgtaagct 1525  
 OY 1143 gtacatgtccaaaagaagaatctccaagaagaacccatcgtgtaagaagaagcccccca 1202  
 Db 1526 ctactacgacagaagaatgcatgacagaagaacccaatgtaagaagtgacattgaa 1585  
 OY 1203 tgcagtgctcaatgagctgtgtgtcttctgatatctctgtgagggccttgaagataag 1262  
 Db 1586 ccccatcttcaatgaaatcttcatcagacacatcccaacttaccctctgctgatacag 1645  
 OY 1263 tgttgaatttgggtttggtatctgaaaggggtccgaatgaggtaaatcgggcaagt 1322  
 Db 1646 catcgaggtctcgttatcgtatcgtatcgtacgacccacaaagtatggtggtgagagct 1705  
 OY 1323 agtcttggtgtgca---gcaagcagaagaactgtgtgaaagacactgcaagaagaatcgtga 1379  
 Db 1706 gatccttgaggacacacagtgltcacagccagtggtgtgtaacactgagagaaggtctgcga 1765  
 OY 1380 ctacccagagagaatgccaagtgcagtgctctgtga 1421  
 Db 1766 gagccccgcgaagcctgtgscgaagtgcacagctctgagcga 1807

RESULT 7  
 AAX41151  
 ID AAX41151 standard; cDNA; 424 BP.  
 AC AAX41151;  
 DT 17-JUN-1999 (first entry)  
 XX Human secreted protein 5' EST SEQ ID NO:95.  
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;  
 KW forensic; gene therapy; chromosome mapping; signal peptide;  
 KW upstream regulatory sequence; cytokine activity; cell proliferation;  
 KW differentiation; haematopoiesis regulation; tissue growth regulation;  
 KW reproductive hormone regulation; chemokine; chemokine; haemostatic;  
 KW thrombolytic; anti-inflammatory; tumour inhibition; ds.  
 XX Homo sapiens.  
 OS WO9906548-A2.  
 XX 11-FEB-1999.  
 PD 31-JUL-1998; 98WO-1801222.  
 PF 01-AUG-1997; 97US-0905135.  
 XX (GEST) GENSET.  
 XX Duclet A, Dumas Milne Edwards J, Lacroix B;  
 PI WPI: 1999-153778/13.  
 DR P-SDB; AAI12318.  
 XX New nucleic acids encoding human secreted proteins - obtained from  
 PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,  
 PT kidney, lung, umbilical cord, placenta and colon tissue  
 XX Claim 1; Page 243; 824pp; English.  
 XX AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for  
 CC human secreted proteins, and encode the proteins given in AAI12261 to.

CC AAI12314, respectively. The proteins given represent the signal peptide  
 CC and an N-terminal fragment of a secreted protein. The nucleic acid  
 CC sequences can be used for producing secreted human gene products. They  
 CC can also be used to develop products for diagnosis and therapy. The  
 CC proteins obtained may have cytokine activity, cell  
 CC proliferation/differentiation activity, haematopoiesis regulating  
 CC activity, tissue growth regulation activity, reproductive hormone  
 CC regulatory activity, chemotactic/chemokinetic activity, haemostatic and  
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter  
 CC sequences. The nucleic acids encoding of a polypeptide or the insertion of a  
 CC polypeptide into a membrane, or importing a polypeptide into a cell.  
 CC  
 XX Sequence 424 BP; 109 A; 112 C; 104 G; 98 T; 1 other;

Query Match 9.38; Score 361; DB 20; Length 424;  
 Best Local Similarity 99.48; Pred No 4.4e-72;  
 Matches 361; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ctcttgcctctccctcgttccagagactggtgacctggcctgtgctgtgttttcag 60  
 Db 62 cctcttgcctctccctcgttccagagactggtgacctggcctgtgctgtgttttcag 121  
 OY 61 cgttcgaaagcggcgcttgagatccaggaagtgaatccagcagcagatttcctt 120  
 Db 122 cgttcgaaagcggcgcttgagatccaggaagtgaatccagcagcagatttcctt 181  
 OY 121 cagcacttcgacagaaacagcagtaaaaaatggtctccatcacacacagcggagaa 180  
 Db 182 cagcacttcgacagaaacagcagtaaaaaatggtctccatcacacacagcggagaa 241  
 OY 181 ttgtatgaatcccccacagtggtggtggtcttcagtgcatttggcctgtgttcacagtc 240  
 Db 242 ttgtatgaatcccccacagtggtggtggtcttcagtgcatttggcctgtgttcacagtc 301  
 OY 241 tctctcttcgatgatctgtgtcagagaaatcatccaagtctaaagaactctcca 300  
 Db 302 tctctcttcgatgatctgtgtcagagaaatcatccaagtctaaagaactctcca 361  
 OY 301 tacaagtgtgcatgcttaagagagttgatatactaccctgaaacctaatagcaaa 360  
 Db 362 tacaagtgtgcatgcttaagagagttgatatactaccctgaaacctaatagcaaa 421  
 OY 361 aag 363  
 Db 422 aag 424

RESULT 8  
 AAT25277  
 ID AAT25277 standard; cDNA to mRNA; 318 BP.  
 AC AAT25277;  
 DT 25-NOV-1996 (first entry)  
 XX Human gene signature H0MG507439.  
 DE Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
 KW human; cloning; mapping; non-biased library; diagnosis; detection;  
 KW cell typing; abnormal cell function; ss.  
 XX Homo sapiens.  
 OS WO9514772-A1.  
 XX 01-JUN-1995.  
 PD 11-NOV-1994; 94WO-JP01916.  
 PF

```

XX 12-NOV-1993; 93JP-0355504.
XX (MATSU/) MATSUBARA K.
XX (OKUBO/) OKUBO K.
XX Matsubara K, Okubo K;
XX WPI: 1995-206931/27.
XX
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX for diagnosis of abnormal cell function, by preparing cDNA that
XX reflects relative abundance of corresp. mRNA in specific human
XX tissues
XX
XX Claim 1: Page 1811: 2245bp; Japanese.
XX
XX A single-stranded DNA (or its complementary strand or the corresp.
XX double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX given in AAT19001-T26837 and which is able to hybridize to part of
XX human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
XX sequences were obtained from 3'-directed cDNA libraries prepared from
XX various human tissues; synthesis of cDNA was initiated from the
XX 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
XX untranslated sequence is unique to a particular mRNA species, almost
XX all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library
XX is constructed so as to reflect accurately the relative abundance of
XX different mRNAs in the particular tissue from which it was derived.
XX The appearance frequency of a given GS in a cDNA library can be
XX determined (esp. using primers and probes derived from the GS
XX sequences) as a means of diagnosing abnormal cell function or for
XX recognising different cell types.
XX
XX Sequence 318 BP; 80 A; 48 C; 66 G; 115 T; 9 other:
XX
XX
XX Query Match
XX Best Local Similarity 94.2%; Score 264.2; DB 16; Length 318;
XX Matches 291; Conservative 0; Mismatches 16; Indels 2; Gaps 2;
XX
XX 3535 gatcctaatacctaattcattcattctctgcaaaatgtaactgtattacaat 3594
XX 1 gatcctaatacctaattcattcattctctgcaaaatgtaactgtattacaat 59
XX
XX 3595 atcaatgtaataatcttgaatcctgttacaacatccgtgcaatgtaataatgtaatt 3654
XX 60 atcaatgtaataatcttgaatcctgttacaacatccgtgcaatgtaataatgtaatt 119
XX
XX 3655 atggttctgtcgttagcacaatctcaacccaatgggaggtatatacatggttgaaga 3714
XX 120 atggttctgtcgttagcacaatctcaacccaatgggaggtatatacatggttgaaga 179
XX
XX 3715 c-gtgaactcggtaattgattgtctcgtatgltgttaactaaatagaagtgtttggaag 3773
XX 180 ctgtgaactcggtaattgattgtctcgtatgltgttaactaaatagaagtgtttggaag 239
XX
XX 3774 gaagcctgtgtgtgtgagcaagtgctgtctcttgtgtgcagcctctgtatgattgttaa 3833
XX 240 gaagcctgtgtgtgtgagcaagtgctgtctcttgtgtgcagcctctgtatgattgttng 299
XX
XX 3834 gaccacgtt 3842
XX 300 gccntgttt 308
XX
XX
XX RESULT 9
XX AAI10603/c
XX ID AAI10603 standard; DNA; 383 BP.
XX AC AAI10603;
XX XX
XX DT 12-OCT-2001 (first entry)
XX XX

```

```

DE Probe #536 for gene expression analysis in human cervical cell sample.
XX
XX Probe: human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn Sq, Hanzel DK, Chen W, Rank DR;
XX
XX WPI: 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25; SEQ ID No 536; 487bp; English.
XX
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 383 BP; 125 A; 71 C; 58 G; 123 T; 6 other:
XX
XX
XX Query Match
XX Best Local Similarity 95.8%; Score 216.6; DB 22; Length 383;
XX Matches 254; Conservative 0; Mismatches 6; Indels 5; Gaps 3;
XX
XX 737 tgaagatcctcccaagagaagaagcataaagtgaactagaatgctgagagaaaactcttg 796
XX 383 TGACGATCCCTCCAGAGAAAGCATTAAGTGAAGCTGAGAAAACCTTGG 324
XX
XX 797 atccagcttctgtgagaaacttaccattctatgtagatcctcaacccaatcc---aag 853
XX 323 ATCCAGCTTTGATGACACCTTTTACATTTCTATGAGATACCTTACACCAATCCNNAGA 264
XX
XX 854 aatggccttgcaacttacaat-cttgaatttgcacaggttttcaaga-gatgatacat 911
XX 263 ATTGNCCTTGACACTTCACATNTTTGAGTTTGACAGGTTTTCAGANGATGATATCAT 204
XX
XX 912 tggggaagtcttaattcctctcgcgggaattgaaatatactgaaagaaaatgttaagaa 971
XX 203 TGGGGGAAGTTCTAATTCCTCTCGGGGAATGGAATTCGAAGGAAAATGTTAATGAA 144
XX
XX 972 tagagatcatcaagagaagaatgtt 996
XX 143 TAGAGAGATCATCAAGAGAATGNT 119
XX
XX
XX RESULT 10
XX AAI1857/c
XX

```

ID AAI31857 standard; DNA; 383 BP.  
 AC AAI31857;  
 XX  
 XX 17-OCT-2001 (first entry)  
 DE Probe #543 used to measure gene expression in human placenta sample.  
 DE Probe: microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder; ss.  
 KW Homo sapiens.  
 OS  
 PN W0200157272-A2.  
 PD 09-AUG-2001.  
 PE 30-JAN-2001; 2001WO-US00663.  
 PF 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI: 2001-488897/53.  
 XX  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 XX Claim 25; SEQ ID No 543; 654bp; English.  
 PS  
 CC The present invention relates to single exon nucleic acid probes (SENPs).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.  
 XX  
 SO Sequence 383 BP; 125 A; 71 C; 58 G; 123 T; 6 other;  
 Query Match 5.6%; Score 216.6; DB 22; Length 383;  
 Best Local Similarity 95.8%; Pred. No. 1.6e-39;  
 Matches 254; Conservative 0; Mismatches 6; Indels 5; Gaps 3;  
 Oy 737 tgaagatccccccagaagaagacataaagtagaagtagctgagaagaaccttg 796  
 Db 383 TGACGATCTCCACAGAGAAGACATAAAGTGAAGTAGAGTCTGAGAAAACCTTG 324  
 Oy atccagcttttgatgagaccttatactctatgagatcccccaaccaatcc---aag 853  
 Db 323 ATCCAGCTTTTGATGAGACCTTATCTATGAGATACCTACACCAATCCNNAGA 264  
 Oy 854 aatgagcttgcacttcaaat-ttgagatttgcagagtttcaaga-gatgatatcat 911  
 Db 263 ATTGGNCTTGCACCTTCAATNTTGAAGTTTGACAGGTTTTCAGAGATGATATCAT 204  
 Oy 912 tggggaagttctaattctctctcgcggaattgaattatctgaagaaatgttaatgaa 971  
 Db 203 TGGGGAAGTTCTAATTCTCTCGGGAATTGAATTATCTGAAGAAATGTTAATGAA 144  
 Oy 972 taagagatcatcaagaagaatgct 996  
 Db 143 TAGAGATCATCAAGAGAAATGNT 119  
 RESULT 11

AAI00535/C  
 ID AAI00535 standard; DNA; 383 BP.  
 AC AAI00535;  
 XX  
 XX 09-OCT-2001 (first entry)  
 DE Probe #526 used to measure gene expression in human breast sample.  
 DE Probe: human; breast disease; breast cancer; development disorder; ss.  
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 KW Homo sapiens.  
 OS  
 PN W0200157270-A2.  
 PD 09-AUG-2001.  
 PE 29-JAN-2001; 2001WO-US00661.  
 PF 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI: 2001-476286/51.  
 XX  
 XX Novel single exon nucleic acid probe used to measuring gene expression  
 PT in a human breast -  
 XX  
 XX Claim 25; SEQ ID No 526; 322bp; English.  
 PS  
 CC The present invention relates to novel single exon nucleic acid probes.  
 CC The present sequence is one such probe. The probes are useful for  
 CC measuring human gene expression in a human breast sample, where the probe  
 CC hybridises at high stringency to a nucleic acid expressed in the human  
 CC breast. The probes are useful for predicting, diagnosing, grading,  
 CC staging, monitoring and prognosing diseases of the human breast,  
 CC particularly those diseases with polygenic aetiology. The diseases  
 CC include: breast cancer, disorders of development, inflammatory diseases  
 CC of the breast, fibrocystic changes, proliferative breast disease and  
 CC non-carcinoma tumours.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SO Sequence 383 BP; 125 A; 71 C; 58 G; 123 T; 6 other;  
 Query Match 5.6%; Score 216.6; DB 22; Length 383;  
 Best Local Similarity 95.8%; Pred. No. 1.6e-39;  
 Matches 254; Conservative 0; Mismatches 6; Indels 5; Gaps 3;  
 Oy 737 tgaagatccccccagaagaagacataaagtagaagtagctgagaagaaccttg 796  
 Db 383 TGACGATCTCCACAGAGAAGACATAAAGTGAAGTAGAGTCTGAGAAAACCTTG 324  
 Oy atccagcttttgatgagaccttatactctatgagatcccccaaccaatcc---aag 853  
 Db 323 ATCCAGCTTTTGATGAGACCTTATCTATGAGATACCTACACCAATCCNNAGA 264  
 Oy 854 aatgagcttgcacttcaaat-ttgagatttgcagagtttcaaga-gatgatatcat 911  
 Db 263 ATTGGNCTTGCACCTTCAATNTTGAAGTTTGACAGGTTTTCAGAGATGATATCAT 204  
 Oy 912 tggggaagttctaattctctcgcggaattgaattatctgaagaaatgttaatgaa 971



```

DB 203 TGGGGAAGTTCTAATCTCTCTCGGGAATGATATATCTGAAGAAATGTTATGAA 144
      |||
QY 972 tagagagatcatcaagaagaatgct 996
      |||
DB 143 TAGAGAGATCATCAAGAGAATGCT 119

```

## RESULT 12

AAF58252/c  
ID AAF58252 standard; DNA; 936 BP.

XX AAF58252;

XX 24-APR-2001 (first entry)

XX Oligonucleotide D1835.

XX Electron-transfer group; ETM; mismatch; genotyping;

XX gene expression; ss.

XX Synthetic.

XX WO200107665-A2.

XX 01-FEB-2001.

XX 26-JUL-2000; 2000WO-US20476.

XX 26-JUL-1999; 99US-0145695.

XX 17-MAR-2000; 2000US-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

XX Umek RM;

XX WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in

XX hybridization assays, e.g. for genotyping, allowing repeat analyses on

XX a single surface -

XX Example 6; Page 127; 159pp; English.

XX The present invention relates to a composition comprising two nucleic

XX acids each containing an electron-transfer group (ETM) having

XX different redox potentials. The invention is used for electronic

XX detection of nucleic acids, especially of substitutions (mismatches)

XX and single-nucleotide polymorphisms, e.g. for genotyping,

XX monitoring gene expression.

XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

XX

XX

XX

XX

XX

XX

XX

XX

QY 1765 agttaaagtattgaatctaaacacgaagatgcttgcttgtaaatctca 1824

DB 548 WWWWWWXX 489

QY 1825 ttatcttcaggctggggaatcaatttcttcttaatcaagatactaaataatgctc 1884

DB 488 WWWWWWXX 429

QY 1885 ctccagtttgatttaatactctgcatgcaatggtgctcgtcatataaagtat 1944

DB 428 WWWWWWXX 369

QY 1945 ctggtcattcagtttggttgtaattatttgatgcaatttatacgaagtaactag 2004

DB 368 WWWWWWXX 309

QY 2005 atctatccaagaagacagctgaacaagctgaagaattatttatacgaagctga 2064

DB 308 WWWWWWXX 249

QY 2065 gaacactggtgctgaataatcttcccccctaaggttacatgtagtcaaat 2124

DB 248 WWWWWWXX 189

QY 2125 tgtaaataataacctacacataagaacacagctgcttgattatctcctgctcacaagc 2184

DB 188 WWWWWWXX 129

QY 2185 ctccagttgctgagaatccctatctaccttgctgaattgtgaattagtagtga 2244

DB 128 WWWWWWXX 69

QY 2245 taaagaataaacttcaactagaatccagttgaagtgcatttctataggaatg 2304

DB 68 WWWWWWXX 9

QY 2305 gtaata 2309

DB 8 WWWWWW 4

## RESULT 13

AAF58254/c  
ID AAF58254 standard; DNA; 936 BP.

XX AAF58254;

XX 24-APR-2001 (first entry)

XX Oligonucleotide D1875.

XX Electron-transfer group; ETM; mismatch; genotyping;

XX gene expression; ss.

XX Synthetic.

XX WO200107665-A2.

XX 26-JUL-2000; 2000WO-US20476.

XX 26-JUL-1999; 99US-0145695.

XX 17-MAR-2000; 2000US-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

XX Umek RM;

XX WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in





Fri Dec 14 10:31:44 2001

us-09-680-121-1.rng

Page 16

Job time: 5761 sec

...

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2001, 10:54:04 ; Search time 130 Seconds  
(without alignments)  
6778.661 Million cell updates/sec

Title: US-09-680-121-1

Perfect score: 3891

Sequence: 1 cctcttgccctccctgctt.....ccccaaaaaaaaaaaaa 3891

Scoring table:

IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 351203 segs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3891	100.0	3891	US-09-036-315-1	Sequence 1, Appl 1
2	51.4	1.3	2599	5266464-1	Patent No. 5266464
3	49.2	1.3	19124	US-08-487-826B-13	Sequence 13, Appl 1
4	47.4	1.2	1207	US-08-872-979-4	Sequence 4, Appl 1
5	47.4	1.2	1214	US-08-817-913-13	Sequence 13, Appl 1
6	47.4	1.2	1232	US-08-817-913-14	Sequence 14, Appl 1
7	47.4	1.2	1352	US-08-817-913-15	Sequence 15, Appl 1
8	47.4	1.2	1734	US-08-817-913-16	Sequence 16, Appl 1
9	47.4	1.2	1920	US-08-817-913-17	Sequence 17, Appl 1
10	46	1.2	1850	US-08-617-860B-32	Sequence 32, Appl 1
11	46	1.2	4098	US-08-605-106-4	Sequence 4, Appl 1
12	44	1.1	1850	US-08-617-860B-32	Sequence 32, Appl 1
13	44	1.1	4098	US-08-605-106-4	Sequence 4, Appl 1
14	44	1.1	7218	US-08-232-463-14	Sequence 14, Appl 1
15	43.6	1.1	1614	US-08-583-672-1	Sequence 1, Appl 1
16	43	1.1	660	US-07-991-867B-32	Sequence 32, Appl 1
17	43	1.1	660	US-08-107-755A-32	Sequence 32, Appl 1
18	43	1.1	660	US-08-544-332-32	Sequence 32, Appl 1
19	43	1.1	1511	US-07-991-867B-8	Sequence 8, Appl 1
20	43	1.1	1511	US-08-107-755A-8	Sequence 8, Appl 1
21	43	1.1	1511	US-08-544-332-8	Sequence 8, Appl 1
22	43	1.1	4810	US-08-852-629-11	Sequence 11, Appl 1
23	43	1.1	4838	US-08-852-629-15	Sequence 15, Appl 1
24	42	1.1	4818	US-08-817-926-27	Sequence 27, Appl 1
25	41.6	1.1	3618	US-07-872-678A-36	Sequence 36, Appl 1
26	41.4	1.1	2799	US-08-212-188-1	Sequence 1, Appl 1
27	41.4	1.1	2799	US-08-970-725-1	Sequence 1, Appl 1

C 28	41.4	1.1	2799	5	PCT-US95-02708-1	Sequence 1, Appl 1
C 29	40.8	1.0	602	1	US-08-764-100-8	Sequence 8, Appl 1
C 30	40.8	1.0	642	1	US-08-764-100-13	Sequence 13, Appl 1
C 31	40.8	1.0	643	1	US-08-764-100-7	Sequence 7, Appl 1
C 32	40.8	1.0	2993	1	US-08-764-100-2	Sequence 2, Appl 1
C 33	40.8	1.0	2993	1	US-08-764-100-10	Sequence 10, Appl 1
C 34	40.8	1.0	3000	1	US-08-764-100-9	Sequence 9, Appl 1
C 35	40.8	1.0	3001	1	US-08-764-100-1	Sequence 1, Appl 1
C 36	40.8	1.0	4673	1	US-07-638-431-1	Sequence 1, Appl 1
C 37	40.8	1.0	4673	5	PCT-US92-00018-1	Sequence 1, Appl 1
C 38	40.8	1.0	7015	4	US-09-177-249-6	Sequence 6, Appl 1
C 39	40.6	1.0	834	4	US-08-998-416-305	Sequence 305, App
C 40	40.6	1.0	5852	1	US-07-867-106-2	Sequence 2, Appl 1
C 41	40.4	1.0	837	4	US-08-998-416-288	Sequence 288, App
C 42	40	1.0	6216	3	US-09-213-053-1	Sequence 1137, App
C 43	39.6	1.0	636	4	US-08-998-416-1137	Sequence 1137, App
C 44	39.6	1.0	3942	3	US-08-480-640A-189	Sequence 189, App
C 45	39.6	1.0	3942	4	US-08-686-968C-189	Sequence 189, App

#### ALIGNMENTS

RESULT 1  
US-09-036-315-1  
; Sequence 1, Application US/09036315  
; Patent No. 6218523  
; GENERAL INFORMATION:  
; APPLICANT: French, Cynthia K.  
; APPLICANT: Schneider, Patrick A.  
; APPLICANT: Yamamoto, Karen K.  
; TITLE OF INVENTION: Prostate Cancer-Specific Marker  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/036,315  
; FILING DATE: 06-MAR-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/047,811  
; FILING DATE: 15-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/041,246  
; FILING DATE: 07-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 018002-000210US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3891 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 151..1425  
; OTHER INFORMATION: /product= "Repro-PC-1.0"



QY 2101 aaggttacatgtagtcaaaattttgtaaaataaaccacacataagaacacatgacctg 2160  
|||  
Db 2101 AAGGTACATGtagtcaaaattttgtaaaataaaccacacataagaacacatgacctg 2160  
QY 2161 gattatcaatcctgctgcacaaagcctcagtgctgcctgagaatccctatgtaacctgtc 2220  
|||  
Db 2161 GATTATTCCTACTCCTGCTCAAGACCCCTCACTGTGGCTGAGAAATCCCTATGTACCTTTGT 2220  
QY 2221 gaaattgtgaaattgtagtgaataaagaataaactcaactgaactgaatccagttaga 2280  
|||  
Db 2221 GAAATGTTGAAATGTAGTGAATGAATAAATAAATTAACCTCACTGAGAAATCCAGTTAGAA 2280  
QY 2281 ggcgaatttctctatagagaatcaggtatagtgctgaagtgtaactttaaagccatcgctt 2340  
|||  
Db 2281 GCGCAATTTCTTATAGGAATAATAGTATAGTGTGCAAGGTGTACTTTTAAGGCCATCGTTT 2340  
QY 2341 gtaccagaagtcgcatgagccaccatagtcctcatttaattatgtgccccagaagaaga 2400  
|||  
Db 2341 GTACCCAGAGTGGCATGGCCACTAAGCTTCATTATTTATTTGTCCTCCAGAAAGA 2400  
QY 2401 ttaaatgtctacttgaagaagactgtgaaagatttttacctgcccagataaagaagtac 2460  
|||  
Db 2401 TTAAGATGCTACTTGAAAGAGCTGTGAAGATTTTACATTCGCCAGATTAAGAGTGTAC 2460  
QY 2461 ttaacaaacaaacaaatgtgaagactacaacaaatcgttcaagcaacttcaataatta 2520  
|||  
Db 2461 TTAACCAACAAACAAATGTGAAGACTACAAATCGTTCAAGACCAATTTCTTAATATATTTA 2520  
QY 2521 catatgtctacgcgaataatctagctgctgcacaaatctagcacaacaaagaagtgttca 2580  
|||  
Db 2521 CATATGTTTACGCCAAATATGTGCTTAGGCTGTCAAAATTTAGCACACAAAGAATGTGTTCA 2580  
QY 2581 ccaatctttctagagctaatgttcttgagcgtgtgtctctatagaagcagttcaagactgtc 2640  
|||  
Db 2581 CATCTTTTCTAGGGCTAAATTTGTGTGAGCTGTGTCTATAGCGAGCTTTACAGACTTGT 2640  
QY 2641 gctctgtatcatcttccagctgcagaggtctcgaatctcagaactcagaacccgttgaattaa 2700  
|||  
Db 2641 GCTCTGTATCATTTTCCAGTGGCCAGGGTCTGAATAATCTTCAAGACCTGTTAGATTAA 2700  
QY 2701 gctgcacccctgcatattctgaagaagaatctagctgagaagtaatgtcactatattgaat 2760  
|||  
Db 2701 GCTGCACCCCTGTGATTTATTTGAAGAATTTAGCTTGAGAGTATGTCACTATATTTGAGT 2760  
QY 2761 tcttagagaagatagagtggaactgtgagctagctgaattatataatgcaagttagaa 2820  
|||  
Db 2761 TCTTAGAGAGATAGTGTGGAACCTTGAGTACAGTTGAATTAATTAATGCAAGTAGAA 2820  
QY 2821 attaaagctactgaaagaatttcaatttgagtcaggttttgctgcagtaacttgaagact 2880  
|||  
Db 2821 ATTAAGTCTACTGAAATAATTTTCAATTTTGTAGTCAAGTGTGTGCTTATAGCAGTT 2880  
QY 2881 tttagaagaatgtgtgatalcacaagtggttgaattcctataaataatgcatlltccaaa 2940  
|||  
Db 2881 TTTGAGAATGTGTGATATCAAGTGTGTGAATAATTCATGAATAATGCATTTTCCAAA 2940  
QY 2941 caactatacatgcttcttataactatgcttaactaaatgaagaagaatglatcaatctctga 3000  
|||  
Db 2941 CAACTATATCAATGCTTTTATGACTATGCTATGTAAGAAAGAAATGATTTACTATCTGTA 3000  
QY 3001 tctacaaagaataaagaataacactcttcttctgctttaaataatgagcttggagtaaaaa 3060  
|||  
Db 3001 TCTACAAAGATTAATAAATACCTCTTTTGTGCTTTAAATGAATTTGGGATTTAAAA 3060  
QY 3061 agcatatttcccaatcatgtctcattccactacaaagtcacccacagacatctgtctc 3120  
|||  
Db 3061 AGCATATTTCCCAATCATGTGTCTATTCCTACTACAAAGTCAACCTGACATCTTGCTC 3120  
QY 3121 caactggcactctctgtgaagaacaaatgaatgaatgaatgaatgaatgaatgaatgaatga 3180  
|||  
Db 3121 CACTGGGCACTCTCTGTGAAGCAACATGAATAATGTAAGTGTAGTGTAGTTGGGGA 3180

QY 3181 agtcaaatgycatatttataatgtagtgcatttgtagtcalatgagccgtgagacagaatata 3240  
|||  
Db 3181 AGTCAAAATGGCCATTTTATGTATGTGCATTTGGTATCATGAGGCGGTGAAGAAATATAT 3240  
QY 3241 gttggaacctgtgaaagaatgttaagggcccaatctaaagtatcttcaacggcagcagaag 3300  
|||  
Db 3241 GTTGGACCTCTGAAAAGTTGTAAAGGGGCCAAATCTAAGATTTCTTCCAGCGCAGGAAG 3300  
QY 3301 ttaagtgttagcagctaggtatgtagttgttgagcgaagccgatttttttttaaa 3360  
|||  
Db 3301 TTAATGTGTAGACACTGAGGTATGTGTGTGAGACAGGCGCATTTTTTTTTAAACA 3360  
QY 3361 tggaaacatgaaacacacacaaacattttaaataatgaatgaatgtttaaataatg 3420  
|||  
Db 3361 TGGAAATGAAGAAACCAACAAACATTTTAAATTAATTAATGATTAATTTGAAATAT 3420  
QY 3421 ttttagctttaaataatgaatgttcttggagtgtgaaagtgtgagtaactatttga 3480  
|||  
Db 3421 TTTTACCTTTTAAATTTTAAAGTGTGTGTGTGAGTGTGAAGGTTGATTAACATATTGCA 3480  
QY 3481 actggtttcagaagaagaagaacaaagaagaatgtgaacagcagggagagctt 3540  
|||  
Db 3481 ACTGTTTTGAGAAAAGAGAAAAGAAACAAAGAAATTTGAACAGGAGGAGATCTT 3540  
QY 3541 aatacctaattcattcattctgcgaatctgactgttcttagaagtgtatcaataatcaat 3600  
|||  
Db 3541 AATACCTAATTTATCATCTTTCTGCAAAATGTCTGTTTGAATGTATATACATATCAAT 3600  
QY 3601 gfgaatacttgaatcgtgttcaaaatcgtcactgtatlaaacaatgtaataatgtt 3660  
|||  
Db 3601 GFGAATATCTGTGAATCTGTACAAATCTGCACGTATTAACATGATTAATTAATTT 3660  
QY 3661 tgtctgattagcacaatctcaccacccaatggygagatataatgtttaaagaactgtga 3720  
|||  
Db 3661 TGCTGATTAAGCAATCTCACCAACCAATGGGAGGTATACATGTTTAAGAACGTGA 3720  
QY 3721 actcggtaattgattgttctgtagtgtgaactcaataagaagtgtttaaagaagcaat 3780  
|||  
Db 3721 ACTCGGTAATTTGATTTGTTCTGATGTGTATACATCAATAGAGTGTGTGGAAGAAAGCT 3780  
QY 3781 ggtgtgtgagacagtgctgtcttcttgcagcagctcgtatagatgttgaagacatg 3840  
|||  
Db 3781 GGTGTGTGAGACAGTGTGTTCTTTTGTGCCAGCTCTGTATGATGTGTGAAGACCATG 3840  
QY 3841 ttgtgaagacagaataaattgctgtcttgcacacacacacacacacacacacacacac 3891  
|||  
Db 3841 TTTGTGAAGACATGAATTAATTTGCTGCTTTGCCCAAAAAA 3891  
  
RESULT 2  
5266464-1  
; Patent No. 5266464  
; APPLICANT: HOUSEY, GERARD  
; TITLE OF INVENTION: METHOD OF SCREENING FOR PROTEIN INHIBITORS  
; AND ACTIVATORS  
; NUMBER OF SEQUENCES: 3  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/392,073  
; FILING DATE: 10-AUG-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 154,206  
; FILING DATE: 10-FEB-1988  
; SEQ ID NO:1:  
; LENGTH: 2599  
5266464-1  
  
Query Match 1.3% Score 51.4; DB 6; Length 2599;  
Best Local Similarity 56.1%; Pred. No. 0.002; 76; Indels 0;  
Matches 97; Conservative 0; Mismatches 0; Gaps 0;  
  
QY 1060 ctactgtgtgtgtctaaagaactcgacatctgcctaaatcgaatgtgtcgcgaacttca 1119  
|||

Fri Dec 14 10:31:47 2001

us-09-680-121-1.rni

Page 4

Db 608 ctcatcgtgtgtgtaagagatgacgtcaaaaatctcgtacctatgaccccaacggcttgcga 667

Oy 1120 gaccccctacgtcaaaagtgaaacccgtgacccctgcgcacaaagaatcctcaagaagaagcct 1179

Oy 668 gactccatagtaataacgttgaacatgataccctgcacccaaaatgtaggaagcagcagac 727

Oy 1180 catgtgaagaataagtcacccccaatgtagtgcataatgacccgtgtgcttga 1232

Oy 728 aagacatctcaaatgtctcccaaccccggtggtgaacgaaacctcaattca 780

RESULT 3  
US-08-487-826B-13/c  
; Sequence 13, Application US/08487826B

```

1  GENERAL INFORMATION:
2  APPLICANT:  Slim, Kim L.
3  APPLICANT:  Chltnis, Chetan
4  APPLICANT:  Miller, Louis H.
5  APPLICANT:  Peterson, David S.
6  APPLICANT:  Su, Xin-zhuan
7  APPLICANT:  Wellens, Thomas E.
8  TITLE OF INVENTION:  BINDING DOMAINS FROM PLASMODIUM VIVAX
9  TITLE OF INVENTION:  AND PLASMODIUM FALCIPARUM ERTHROCYTE BINDING PROTEINS
10 NUMBER OF SEQUENCES:  45
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE:  Knobble Martens Olson & Bear
13 STREET:  620 Newport Center Drive  16th Floor
14 CITY:  Newport Beach
15 STATE:  California
16 COUNTRY:  US
17 ZIP:  92660
18
19 COMPUTER READABLE FORM:
20 MEDIUM TYPE:  Floppy disk
21 COMPUTER:  IBM PC Compatible
22 OPERATING SYSTEM:  PC-DOS/MS-DOS
23 SOFTWARE:  PatentIn Release #1.0, Version #1.25
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER:  US-/08/487, 826B
26 FILING DATE:  10-SEP-1993
27 CLASSIFICATION:  435
28 ATTORNEY/AGENT INFORMATION:
29 NAME:  Israelisen, Ned
30 REGISTRATION NUMBER:  29,655
31 REFERENCE/DOCKET NUMBER:  NIH121.001CPI
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE:  (619) 235-8550
34 TELEFAX:  (619) 235-0176
35 INFORMATION FOR SEQ ID NO:  13:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH:  19124 base pairs
38 TYPE:  nucleic acid
39 STRANDEDNESS:  single
40 TOPOLOGY:  linear
41 MOLECULE TYPE:  CDNA
42 HYPOTHETICAL:  NO
43 ANTI-SENSE:  NO
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
9
```

Query Match	Similarity	1.38;	Score 49.2;	DB 2;	Length 19124;
Best Local	Similarity	50.78;	Pred. No. 0.018;		
Matches 144;	Conservative	0;	Mismatches 138;	Indels 2;	Gaps 1.
QY 3409	tttgcgaatagttttagctttaaataaactgaagtgcttttgagtgtaaaagtgtgata	3468			
Db 449	TTATTATATTTATAATAATTTTTTTTTTTTTTTTTTTTTTTTGTGATGCTGAA	350			
QY 3469	aaactatttgcgaactgttttcgaaagaagaaagaaacaaacaaagaattgaacag	3528			
Db 389	TTACTTTTGTGTTTGTTCGTGATAAAGAAATTTAGAAATTTAAAGAAAGCAATTAAGATTAA	350			
QY 3529	caggagagacttaataactaatccaattcgcgaagaatgacgtgtttagaatgat	3588			

Db 329 TAAAAATTAAGAAATTGCTATATTTATTTAAATAACAGAAATTTATTTGATTTTAT 270  
Qy 3589 taacatac-atgtgatatcttgaaatccctgtacacatccgcactgtataaacat 364  
Db 269 TAATTTATTTATTAAGAAACTATATAATTTGTATTTAAATAATTTTATATGAAATT 210  
Qy 3647 gtaataaattgttcgtctgttgccaactcacaccacat 3690  
Db 209 ATATATTTTTCCTCGATGATTTTCAATTTATTAACAAAAT 166

Db  
209 ATATATTTTTTTTCGCTCGGATATTCAAATTAAATAACAAAT 166

RESULT 4  
US-08-872-979-4  
; Sequence 4, Application US/08872979

```

GENERAL INFORMATION:
APPLICANT: Billman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Cortley, Nell C.
TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872, 979
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0320 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1207 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGNOF12
CLONE: 1003941

```

Query Match	Similarity	1.2%	Score 47.4	DB 3	Length 1207
Best Local	Similarity 47.58		Pred. No. 0.014	Mismatches 191	Indels 12
Matches 184	Conservative	0			Gaps 1
QY	1015	ggtgagttactgacatctctctctgctatcagtcacccaacacctaactggtgtgttc	1074		
Db	697	GGGAGAGCTGCTCTCTCTCTCCGGTACCTGCCAGACAGCGCGGTGACGGTGTGTGTG	756		
QY	1075	ttaaaagctgcacatctgcgctaatactgatgtgtccgagacttgcagatccctatgtcaaa	1134		
b	757	CTGGAGAGCTGCAGAGCTGCGGT-----CCAGAGATTTCAGAGACCCAGAGGTAAG	804		



```

; INFORMATION FOR SEQ ID NO: 14:
;
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 1232 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-817-913-14

Query Match 1.28; Score 47.4; DB 4; Length 1232;  
 Best Local Similarity 46.58; Pred. No. 0.015;  
 Matches 193; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

QY 2427 aagatttttaccatccagataaaagtgttacttaaccacaacaatgtaagacta 2486  
 DB 449 ATGCATCTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 390  
 QY 2487 caaatcgttcaagagcaattcctaataattacataatgltcagcaaatatgcttag 2546  
 DB 389 CATCAAAATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 330  
 QY 2547 gctgtcaaatltagcacacaagaatglttccatccttctcaggctaattgctt 2606  
 DB 329 TTCATTTATTTACAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 274  
 QY 2607 gaggctgttctatagagcagtttacagactgtgtctgtatcatlctccagtgccagg 2666  
 DB 275 GAATCATTAATTTCTAAATAATTAATTTTATTAATTAATTAATTAATTAATTAATTA 216  
 QY 2667 gtctgaattcatltagaacctgttagatlaaagctgcagccctgtgattatcgaaga 2726  
 DB 215 ATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 156  
 QY 2727 aatgaagcttgagaglaatgtccatatttgaagctttagagagagtggtggaactg 2786  
 DB 155 AAATTAATTTCAAAATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 96  
 QY 2787 agtcaagcttgaattatlaaatatgcgaagctagaatgaagctcactgaagaattt 2841  
 DB 95 AACATCATGATTAATTAATTAATTAATTTGCTTACATATTTTAAAGAAATGT 41

RESULT 7

US-08-817-913-15/C  
 Sequence 15, Application US/08817913  
 Patent No. 6184443  
 GENERAL INFORMATION:  
 APPLICANT: Pedersen, Rolf  
 APPLICANT: Lund, Marianne  
 APPLICANT: Okkels, Finn  
 APPLICANT: Kreiberg, Jette  
 TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO  
 NUMBER OF SEQUENCES: 27  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Knobbe, Martens, Olson & Bear  
 STREET: 620 Newport Center Drive 16th Floor  
 CITY: Newport Beach  
 STATE: CA  
 COUNTRY: U.S.A.  
 ZIP: 92660  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTEO Version 1.5  
 CURRENT APPLICATION DATA: US/08/817,913  
 APPLICATION NUMBER: 800  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/EP95/02196  
 FILING DATE: 06-JUN-1995  
 APPLICATION NUMBER: GB941286.7  
 FILING DATE: 21-OCT-1994

ATTORNEY/AGENT INFORMATION:  
 NAME: Altman, Daniel E.  
 REGISTRATION NUMBER: 34,115  
 REFERENCE/DOCKET NUMBER: DYO10.001APC  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 714-760-0404  
 TELEFAX: 714-760-9502  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1352 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-817-913-15

Query Match 1.28; Score 47.4; DB 4; Length 1352;  
 Best Local Similarity 46.58; Pred. No. 0.015;  
 Matches 193; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

QY 2427 aagatttttaccatccagataaaagtgttacttaaccacaacaatgtaagacta 2486  
 DB 569 ATGCATCTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 510  
 QY 2487 caaatcgttcaagagcaattcctaataattacataatgltcagcaaatatgcttag 2546  
 DB 509 CATCAAAATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 450  
 QY 2547 gctgtcaaatltagcacacaagaatglttccatccttctcaggctaattgctt 2606  
 DB 449 TTCATTTATTTACAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 394  
 QY 2607 gaggctgttctatagagcagtttacagactgtgtctgtatcatlctccagtgccagg 2666  
 DB 395 GAATCATTAATTTCTAAATAATTAATTTTATTAATTAATTAATTAATTAATTAATTA 336  
 QY 2667 gtctgaattcatltagaacctgttagatlaaagctgcagccctgtgattatcgaaga 2726  
 DB 335 ATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 276  
 QY 2727 aatgaagcttgagaglaatgtccatatttgaagctttagagagagtggtggaactg 2786  
 DB 275 AAATTAATTTCAAAATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 216  
 QY 2787 agtcaagcttgaattatlaaatatgcgaagctagaatgaagctcactgaagaattt 2841  
 DB 215 AACATCATGATTAATTAATTAATTAATTTGCTTACATATTTTAAAGAAATGT 161

RESULT 8

US-08-817-913-16/C  
 Sequence 16, Application US/08817913  
 Patent No. 6184443  
 GENERAL INFORMATION:  
 APPLICANT: Pedersen, Rolf  
 APPLICANT: Lund, Marianne  
 APPLICANT: Okkels, Finn  
 APPLICANT: Kreiberg, Jette  
 TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO  
 NUMBER OF SEQUENCES: 27  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Knobbe, Martens, Olson & Bear  
 STREET: 620 Newport Center Drive 16th Floor  
 CITY: Newport Beach  
 STATE: CA  
 COUNTRY: U.S.A.  
 ZIP: 92660  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 OPERATING SYSTEM: DOS

```

: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/817,913
: FILING DATE: 15-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP95/02196
: FILING DATE: 06-JUN-1995
: APPLICATION NUMBER: GB941286.7
: FILING DATE: 21-OCT-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Altman, Daniel E
: REGISTRATION NUMBER: 34,115
: REFERENCE/DOCKET NUMBER: DYOUI0.001APC
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 714-760-0404
: TELEFAX: 714-760-9502
: TELEX:
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1734 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-817-913-16

```

```

Query Match      1.2%; Score 47.4; DB 4; Length 1734;
Best Local Similarity 46.5%; Pred. No. 0.017;
Matches 193; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

Oy 2427 aagatttttccatgcccagataaaagtgcttacttaaccacaacaagaatglaagacta 2486
Db 768 ATGATCTTTAAATTTTAAATATAGTATATATATATATATATATATATATATATATAT 709

Oy 2487 caaaatcgltcaagagcaattcctaataattacataatgltcagcacaatatgcttag 2546
Db 708 CATCAAAATTTGAAATATATTAACAACATGATTAATGAACATCATAGAAATTAACCTCG 649

Oy 2547 gctgccaattagcacacaagaatggttccactactcttctcaggactaattgctct 2606
Db 648 TTCATTATTATACAAAGCATTAATATCTTAATCTTAATCTTAATCTTAATCTTAATCTTA 593

Oy 2607 gagcgtgtcctatagagcaggttccagactggtgtctgtatcatcttccagtgccag 2666
Db 594 GAATCATTTATTTCTAAAAATATATTTTATTTATTTATTTATTTATTTATTTATTTATTT 535

Oy 2667 gtctgaaatcattcagaacctgttagatlaaagctgcacctgtgattatltgaagaag 2726
Db 534 ATAATAAATTAACCTTACTAGATATATATAAAGCATCAGCATGATGATGATGATGATGAT 475

Oy 2727 aattagcttgagagtaatgltcactatattgagttccttagagaagtaatgagtggaact 2786
Db 474 AAATTAATTTTCAAAATAGCTATATATAAATATGATTAATCTTTGAGATTTATGACAAACAT 415

Oy 2787 agtacagtgaaattataataatagcaagttgaattgaagttactgtaaaattt 2841
Db 414 AAGATCAATGATTAAGTTAAAGTAATTTGCTTACATATTTTAAAGAAATGT 360

```

```

RESULT 9
US-08-817-913-17/C
: Sequence 17, Application US/08817913
: Patent No. 618443
: GENERAL INFORMATION:
: APPLICANT: Pedersen, Rolf
: APPLICANT: Lund, Marianne
: APPLICANT: Okels, Finn
: APPLICANT: Kreiberg, Jette
: TITLE OF INVENTION: PROMOTER SEQUENCE FROM POTATO
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:

```

```

: ADDRESSEE: Knobbe, Martens, Olson & Bear
: STREET: 620 Newport Center Drive 16th Floor
: CITY: Newport Beach
: STATE: CA
: COUNTRY: U.S.A.
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/817,913
: FILING DATE: 15-SEP-1997
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/EP95/02196
: FILING DATE: 06-JUN-1995
: APPLICATION NUMBER: GB941286.7
: FILING DATE: 21-OCT-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Altman, Daniel E
: REGISTRATION NUMBER: 34,115
: REFERENCE/DOCKET NUMBER: DYOUI0.001APC
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 714-760-0404
: TELEFAX: 714-760-9502
: TELEX:
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1920 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-817-913-17

```

```

Query Match      1.2%; Score 47.4; DB 4; Length 1920;
Best Local Similarity 46.5%; Pred. No. 0.018;
Matches 193; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

Oy 2427 aagatttttccatgcccagataaaagtgcttacttaaccacaacaagaatglaagacta 2486
Db 768 ATGATCTTTAAATTTTAAATATAGTATATATATATATATATATATATATATATATAT 709

Oy 2487 caaaatcgltcaagagcaattcctaataattacataatgltcagcacaatatgcttag 2546
Db 708 CATCAAAATTTGAAATATATTAACAACATGATTAATGAACATCATAGAAATTAACCTCG 649

Oy 2547 gctgccaattagcacacaagaatggttccactactcttctcaggactaattgctct 2606
Db 648 TTCATTATTATACAAAGCATTAATATCTTAATCTTAATCTTAATCTTAATCTTAATCTTA 593

Oy 2607 gagcgtgtcctatagagcaggttccagactggtgtctgtatcatcttccagtgccag 2666
Db 594 GAATCATTTATTTCTAAAAATATATTTTATTTATTTATTTATTTATTTATTTATTTATTT 535

Oy 2667 gtctgaaatcattcagaacctgttagatlaaagctgcacctgtgattatltgaagaag 2726
Db 534 ATAATAAATTAACCTTACTAGATATATATAAAGCATCAGCATGATGATGATGATGATGAT 475

Oy 2727 aattagcttgagagtaatgltcactatattgagttccttagagaagtaatgagtggaact 2786
Db 474 AAATTAATTTTCAAAATAGCTATATATAAATATGATTAATCTTTGAGATTTATGACAAACAT 415

Oy 2787 agtacagtgaaattataataatagcaagttgaattgaagttactgtaaaattt 2841
Db 414 AAGATCAATGATTAAGTTAAAGTAATTTGCTTACATATTTTAAAGAAATGT 360

```

```

RESULT 10
US-08-617-860B-32

```

Sequence 32, Application US/08617860B  
Patent No. 6133506  
GENERAL INFORMATION:  
APPLICANT: Topfer, R., Baulor, J., Bothmann, H., Filasak, E.,  
APPLICANT: Hivick-Grandpierre, C., Klein, B., Martini, N.,  
APPLICANT: M. Iler, A., Schulte, W., Voeltz, M., Walek, J.,  
APPLICANT: Scheil, J.  
TITLE OF INVENTION: Promoters  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Steinberg, Raskin & Davidson, P.C.  
STREET: 1140 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/617,860B  
FILING DATE: 01-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP94/02950  
FILING DATE: 05-SEP-1994  
APPLICATION NUMBER: DE P4329951.2  
FILING DATE: 04-SEP-1993  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1850 Base pairs  
TYPE: Nucleic acid  
STRANDEDNESS: Double stranded  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Cuphea lanceolata  
IMMEDIATE SOURCE:  
LIBRARY: genomic Lambda FIX II  
CLONE: C1Reg1  
FEATURE:  
NAME/KEY: CAAT-Signal  
LOCATION: 1428..1432  
FEATURE:  
NAME/KEY: TATA-Signal  
LOCATION: 1553..1556  
FEATURE:  
NAME/KEY: Transcription start  
LOCATION: 1585  
FEATURE:  
NAME/KEY: Legumindox  
LOCATION: 1642..1657  
FEATURE:  
NAME/KEY: Startcodon  
LOCATION: 1797..1799  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1797..1850  
US-08-617-860B-32

Query Match 1.28: Score 46; DB 3; Length 1850;  
Best Local Similarity 49.28; Pred. NO. 0.039; Indels 0; Gaps 0;  
Matches 121; Conservative 0; Mismatches 125;  
DB 2824 aagctcactgaaattacatttgagtcaggtttgtgctcagtaactagcagtttt 2883  
147 AATTTCGCGGATATATTTTGAATTTTATGATTTATTTGAATTTTGAATTTTGAATTTTGA 206  
QY 2884 gagaatgctgtgatatacagctgttctgtaaatctatgaaataagcatttcacaaca 2943

DB 207 AATATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 266  
QY 2944 ctatcacgtctttttagctactgacctgaatgaaagaatgtatcattcgtatgt 3003  
DB 267 TTTAAATTTAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 326  
QY 3004 acaagattaaataacaccccttttggctttaaataagcttgagattaaaaagc 3063  
DB 327 AATGTTGAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 386  
QY 3064 atattc 3069  
DB 387 ATTTT 392  
RESULT 11  
US-08-605-106-4  
Sequence 4, Application US/08605106  
Patent No. 5910631  
GENERAL INFORMATION:  
APPLICANT: Topfer, R.  
APPLICANT: Martini, N.  
APPLICANT: Scheil, J.  
TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
STREET: P.O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,106  
FILING DATE: 23-SEPT-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP94/02935  
FILING DATE: 01-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Woessner, Warren D  
REGISTRATION NUMBER: 30,440  
REFERENCE/DOCKET NUMBER: 235,001US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-373-6900  
TELEFAX: 612-339-3061  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4098 Base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double stranded  
TOPOLOGY: linear  
MOLECULE TYPE: DNS (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Cuphea lanceolata  
IMMEDIATE SOURCE:  
LIBRARY: genomic Lambda FIX II  
CLONE: C1Reg1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3132  
LOCATION: ..3303, 3391..3459, 3672..3941)  
FEATURE:  
NAME/KEY: Startcodon

```
LOCATION: 1797..1799
FEATURE:
NAME/KEY: exon II
LOCATION: 1787..2294
FEATURE:
NAME/KEY: intron II
LOCATION: 2295..2657
FEATURE:
NAME/KEY: exon III
LOCATION: 2658..2791
FEATURE:
NAME/KEY: intron III
LOCATION: 2792..2897
FEATURE:
NAME/KEY: exon IV
LOCATION: 2898..3011
FEATURE:
NAME/KEY: intron IV
LOCATION: 3012..3131
FEATURE:
NAME/KEY: exon V
LOCATION: 3132..3303
FEATURE:
NAME/KEY: intron V
LOCATION: 3304..3390
FEATURE:
NAME/KEY: exon VI
LOCATION: 3391..3459
FEATURE:
NAME/KEY: intron VI
LOCATION: 3460..3671
FEATURE:
NAME/KEY: exon VII
LOCATION: 3672..3941
FEATURE:
NAME/KEY: Stopcodon
LOCATION: 3942..3944
US-08-605-106-4
```

Query Match 1.2%; Score 46; DB 2; Length 4098;  
Best Local Similarity 49.2%; Pred. No. 0.056;  
Matches 121; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

```
QY 2824 aagtcacgaaattacatttgagcagcttctgctgacgactttagcagctttt 2883
DB 147 AATTTGGGGATATATTTGTAATTTTATGAAATTTTGAATTTTGAATTTTAA 206
QY 2884 gagaatgcttgatatacacagtgcttgaattctatgaaatgcatccaaaca 2943
DB 207 AATATTTTAAATTTAAATATATTTAAATCTTTTAAAAAAATTTTAAATAT 266
QY 2944 ctatacatgcttcttactatgactacgctaactgaaagaatgcatctctgact 3003
DB 267 TATATAATTTAGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 326
QY 3004 acaagattaaataacacactcttcttctgcttcaaatgacttgaggaataaagc 3063
DB 327 AATATGTTGAATTTTAAATATTTTGGTTTAAATAATATTTTAAAGTTTAAAT 386
QY 3064 atattt 3069
DB 387 ATTTT 392
```

US-08-617-860B-32/C  
Sequence 32, Application US/08617860B  
Patent No. 6133506  
GENERAL INFORMATION:  
APPLICANT: Typfer, R., Bautor, J., Bothmann, H., Filasak, E.,  
APPLICANT: Hviricko-Grandpierre, C., Klein, B., Martini, N.,  
APPLICANT: M. Iler, A., Schulte, W., Voeltz, M., Malek, J.,

```
APPLICANT: Schell, J.  
TITLE OF INVENTION: Promoters  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Steinberg, Raskin & Davidson, P.C.  
STREET: 1140 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/617,860B  
FILING DATE: 01-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP94/02950  
FILING DATE: 05-SEP-1994  
APPLICATION NUMBER: DE P432951.2  
FILING DATE: 04-SEP-1993  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1850 base pairs  
TYPE: Nucleic acid  
STRANDEDNESS: Double stranded  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Cuphea lanceolata  
IMMEDIATE SOURCE:  
LIBRARY: genomic Lambda FIX II  
CLONE: CITEg1  
FEATURE:  
NAME/KEY: CAAT-Signal  
LOCATION: 1428..1432  
FEATURE:  
NAME/KEY: TATA-Signal  
LOCATION: 1553..1556  
FEATURE:  
NAME/KEY: Transcription start  
LOCATION: 1585  
FEATURE:  
NAME/KEY: Leguminox  
LOCATION: 1642..1657  
FEATURE:  
NAME/KEY: Startcodon  
LOCATION: 1797..1799  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1797..1850  
US-08-617-860B-32
```

Query Match 1.1%; Score 44; DB 3; Length 1850;  
Best Local Similarity 56.1%; Pred. No. 0.12;  
Matches 83; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```
QY 3344 atttttttttttaacatggaacatgaaacacaaacacattttaaataatg 3403
DB 370 ATATATTTTAAACCAAAAATTTTAAATTCACAAATTTTAAATAAAAAA 311
QY 3404 gataatttgtaagtgtttagctttaaataatgaagtgttgagtggaagtt 3463
DB 310 ACTAATTTTAAATATTTTAAATAATTTTAAACCTAATTTTAAATAATTTT 251
QY 3464 gagtaaacatttgcactggtttca 3491
DB 250 TTTTAAAGAAATTTAAATATTTTAA 223
```



```

; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)683-4109
; TELEFAX: (703)683-4109
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PT2gpt-F15
; US-08-232-463-14

```

```

Query Match          1.1%; Score 44; DB 1; Length 7218;
Best Local Similarity 3.3%; Pred. No. 0.23; Mismatches 135; Indels 0; Gaps 0;
Matches 11; Conservative 190;

```

```

QY 3201 tatgtcattgtatcatgagcgctggaacagaaatatagttgacccctgaagaattg 3260
   || ||||| ||||| : : : : : : : : : : : : : : : : : : : : : : :
DB 1450 TAGAGGAATTGTGTCRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1391

QY 3261 taaggggcaaatctaagtattcttcacggcagccagaagttaattgtgtgagctga 3320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1390 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1331

QY 3321 ggtatgtgtgtgagcagcgccgattttttttaaatgaaacaaagaaaccacaa 3380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1330 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1271

QY 3381 caaacattttaaatlaaatgataattgtlaaagtatttaactttaaatlaa 3440
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1270 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1211

QY 3441 agtcttttgagtgtaaaagtgtgaaactatttgaactgttccagaagaaga 3500
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1210 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1151

QY 3501 aaagaacacaaagaattgaaacagcgagggaga 3536
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1150 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

```

```

RESULT 15
US-08-583-672-1/C
; Sequence 1, Application US/08583672
; Patent No. 5741673
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; APPLICANT: Leonard, James N.
; TITLE OF INVENTION: A NOVEL HOMEOBOX FACTOR THAT STIMULATES
; TITLE OF INVENTION: INSULIN EXPRESSION IN PANCREATIC ISLET CELLS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,672

```

```

; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/106,936
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9422
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1614 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 331..1182
; OTHER INFORMATION: /product="TFP-1 Homeobox-type
; OTHER INFORMATION: transcription factor"
; US-08-583-672-1

```

```

Query Match          1.1%; Score 43.6; DB 1; Length 1614;
Best Local Similarity 80.8%; Pred. No. 0.15; Mismatches 14; Indels 1; Gaps 1;
Matches 63; Conservative 0;

```

```

QY 105 caggcagtttcccttcacacactcgagacagacagcagtaaaaaatgtccgcac 164
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 78 CTGCTGGAATTCTTTCACACACTCGAGACAGACGACGAG-AAACATGCTCTATCAC 20

QY 165 caccagccggagaatt 182
   ||||| |||||
DB 19 CACACACCGCGTGAATT 2

```


Search completed: December 12, 2001, 12:24:59  
Job time: 5455 sec

Fri Dec 14 10:31:47 2001

us-09-680-121-1.rni

---

Page 12





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2001, 10:53:49 ; Search time 24.92 Seconds  
(without alignments)  
1263.288 Million cell updates/sec

Title: US-09-680-121-2

Perfect score: 2201  
Sequence: 1 MADITTSREFFDEIPYVGI.....KEICDYPRQIAKMHVLCDG 425

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: A.GeneSeq\_1101.\*  
2: /SID88/gcgdata/geneSeq/geneSeq/AA1980.DAT.\*  
3: /SID88/gcgdata/geneSeq/geneSeq/AA1981.DAT.\*  
4: /SID88/gcgdata/geneSeq/geneSeq/AA1982.DAT.\*  
5: /SID88/gcgdata/geneSeq/geneSeq/AA1983.DAT.\*  
6: /SID88/gcgdata/geneSeq/geneSeq/AA1984.DAT.\*  
7: /SID88/gcgdata/geneSeq/geneSeq/AA1985.DAT.\*  
8: /SID88/gcgdata/geneSeq/geneSeq/AA1986.DAT.\*  
9: /SID88/gcgdata/geneSeq/geneSeq/AA1987.DAT.\*  
10: /SID88/gcgdata/geneSeq/geneSeq/AA1988.DAT.\*  
11: /SID88/gcgdata/geneSeq/geneSeq/AA1989.DAT.\*  
12: /SID88/gcgdata/geneSeq/geneSeq/AA1990.DAT.\*  
13: /SID88/gcgdata/geneSeq/geneSeq/AA1991.DAT.\*  
14: /SID88/gcgdata/geneSeq/geneSeq/AA1992.DAT.\*  
15: /SID88/gcgdata/geneSeq/geneSeq/AA1993.DAT.\*  
16: /SID88/gcgdata/geneSeq/geneSeq/AA1994.DAT.\*  
17: /SID88/gcgdata/geneSeq/geneSeq/AA1995.DAT.\*  
18: /SID88/gcgdata/geneSeq/geneSeq/AA1996.DAT.\*  
19: /SID88/gcgdata/geneSeq/geneSeq/AA1997.DAT.\*  
20: /SID88/gcgdata/geneSeq/geneSeq/AA1998.DAT.\*  
21: /SID88/gcgdata/geneSeq/geneSeq/AA1999.DAT.\*  
22: /SID88/gcgdata/geneSeq/geneSeq/AA2000.DAT.\*  
23: /SID88/gcgdata/geneSeq/geneSeq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2201	100.0	425	19	AAW74584
2	2201	100.0	425	19	AAW75782
3	1162	52.8	431	22	AAW39577
4	1159	52.7	431	22	AAW39577
5	1154	52.4	486	22	AAW41363
6	603.5	27.4	422	17	AAW77222
7	464.5	21.1	474	21	AAW41973
8	464.5	21.1	474	22	AAW40253
9	464.5	21.1	507	22	AAW42039
10	428	19.4	412	18	AAW25032
11	408	18.5	704	15	AAW57421

12	393.5	17.9	405	20	AAW27275
13	384.5	17.5	400	17	AAW01114
14	384.5	17.5	400	20	AAW34429
15	373	16.9	72	20	AAW12318
16	370	16.8	456	22	AAW35777
17	348.5	15.8	405	22	AAW30777
18	338.5	15.4	375	20	AAW67702
19	319.5	14.5	219	21	AAW56870
20	310	14.1	60	19	AAW30696
21	298	13.5	426	22	AAW2056
22	267	12.1	132	22	AAW05973
23	253	11.5	123	22	AAW05976
24	235.5	10.7	336	22	AAW04779
25	226	10.3	46	19	AAW30697
26	225	10.2	101	21	AAW54122
27	215	9.8	550	22	AAW94880
28	213.5	9.7	228	21	AAW02306
29	202	9.2	490	22	AAW04764
30	181.5	8.2	293	19	AAW63683
31	181.5	8.2	358	22	AAW32692
32	180	8.2	34	19	AAW30694
33	177.5	8.1	696	21	AAW30694
34	177.5	8.1	697	9	AAW82018
35	177	8.0	313	22	AAW04760
36	176	8.0	319	22	AAW41888
37	173.5	7.9	834	18	AAW24227
38	171.5	7.8	280	21	AAW19565
39	169	7.7	33	19	AAW30698
40	169	7.7	313	22	AAW40102
41	166	7.5	31	19	AAW30693
42	164	7.5	672	21	AAW94259
43	163	7.4	672	17	AAW94765
44	161	7.3	672	16	AAW66726
45	161	7.3	1104	22	AAW94266

## ALIGNMENTS

RESULT 1	
AAW74584	standard; Protein; 425 AA.
ID	
XX	
AC	AAW74584;
XX	
DT	11-JAN-1999 (first entry)
XX	
DE	Repro-PC-1.0 prostate cancer-specific marker.
XX	
KW	Repro-PC-1.0; prostate cancer; marker; synaptotagmin; human; diagnosis; vaccine; therapy.
XX	
OS	Homo sapiens.
XX	
PN	W09839447-A1.
XX	
PD	11-SEP-1998.
XX	
PF	06-MAR-1998; 98W0-15374488.
XX	
PR	15-MAY-1997; 97US-0047811.
PR	07-MAR-1997; 97US-0041246.
XX	
PA	(REPR-) REPROGEN INC.
XX	
PI	French CK, Schneider PA, Yamamoto KK;
XX	
DR	WPI: 1998-506363/43.
XX	
DR	N-PSDB: AAV54208.
XX	
PT	New isolated Repro-PC-1.0 polynucleotide(s) - are used to develop products for the detection and prophylactic and therapeutic treatment of prostate cancer
PT	

Mouse Doc2alpha pr  
Brain-specific pro  
Human Doc2-alpha.  
Human 5' EST seque  
Human protein sequ  
Human shear stress  
A human membrane f  
Human prostate can  
Repro-PC-1.0 immu  
Human neuronal apo  
Rat synaptotagmin  
Rat synaptotagmin  
Human vesicle traf  
Repro-PC-1.0 immu  
Human pancreatic c  
Human protein sequ  
Human secreted pro  
Human vesicle traf  
Human secreted pro  
Human protein sequ  
Repro-PC-1.0 immu  
Rat protein kinase  
Protein kinase C-I  
Human vesicle traf  
Human vesicle traf  
Active type R-Ras  
Rat mutated protei  
Repro-PC-1.0 immu  
Human polypeptide  
Repro-PC-1.0 immu  
cDNA encoding prot  
Type III (alpha-ty  
protein kinase C m  
Human protein sequ

XX Claim 1; Page 68-69; 87pp; English.

PS Repro-PC-1.0 is a novel prostate cancer-specific marker, which  
 XX Represents a novel human brain synaptotagmin isoform that may  
 CC function in exocytosis and endocytosis pathways. Its amino acid  
 CC sequence was deduced from cDNA clones (See AAV54206) isolated from  
 CC a male LNCaP tumour cDNA library. The invention provides  
 CC Repro-PC-1.0 polypeptides, fragments and immunogenic peptides  
 CC (see also AAV30693-98) capable of inducing a Class I MHC-restricted  
 CC cytotoxic T-lymphocyte response or a Class II MHC-restricted  
 CC immune response against cells expressing Repro-PC-1.0. Also  
 CC claimed are Repro-PC-1.0 polypeptide analogues that act as decoys,  
 CC antibodies specific for Repro-PC-1.0, and methods for detecting  
 CC Repro-PC-1.0 polypeptide in a sample, for diagnosis of prostate  
 CC cancer (by detecting Repro-PC-1.0 mRNA or polypeptide in a sample),  
 CC for detecting prostate cancer cells (e.g. using the antibodies of  
 CC the invention), for following the progress of prostate cancer, for  
 CC treatment of prostate cancer (e.g. using a Repro-PC-1.0 decoy  
 CC polypeptide), and a polypeptide vaccine for eliciting an immune  
 CC response against Repro-PC-1.0.

XX Sequence 425 AA:

Query Match 100.0%; Score 2201; DB 19; Length 425;

Best Local Similarity 100.0%; Pred. No. 2, 8e-196;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPITTSREEDPEIPVYVIGFSAFGLVTVSLFAMICCORSSKSNKTPPYKRVHVLKGV 60  
 DB 1 mapitstreedeipvtvgifsafiglvfsvlfawiccgqksksnktppykrtvhlkgy 60  
 QY 61 DIYPENLNKKKFGADDNENKRPAPVKNLSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120  
 DB 61 diypenlnskkfgaddnknknpavpknslhldlekrdngnfpktnlkpspsdlen 120  
 QY 121 ATPKFLGEGKESVPSLSKSTSLTSEEKOEKLTLPFSLFENFEKRAFYVNINKEARGL 180  
 DB 121 atpkflgegesvpslskstsltseekoeikltlpfslfeynfekafvnlkeargl 180  
 QY 181 PAMDEOSMTSDPYIKMTILPEKKHKVKTIVLAKTIDPAFDETFYFGIPYQIOELAHF 240  
 DB 181 pamdeqsmtsdpyikmtllpekhhkvktvrlktdpafdefctfygipqyqelahlf 240  
 QY 241 TILSFDRSRDDIIGEVLIPLSGIELSEKMLMNRDIIRKRVKSSGREGLLISLCYOST 300  
 DB 241 tilsfdrsrddiigevliplsgielseggkmlmnrailkrvrkssgrygelllslcyost 300  
 QY 301 TNLTLVVVNLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNAVFNELEVF 360  
 DB 301 tnltlvvvnlkarhlpkssdvsglsdpvykvnlyhakkriskkthvkctpnavfnelvfe 360  
 QY 361 DIPCEGLDIEVEFLVLDSEKSRNEVIGOLVGAAGEGTGEHMKKEICDPPRQIAKWH 420  
 DB 361 dipcegladiesveflvldsergsrnevigolvgaagegtgehmkkeicdyprtrjakwh 420  
 QY 421 VIICDG 425  
 DB 421 viicdg 425

RESULT 2

ID AAV5782 standard; Protein; 425 AA.

AAV5782;

21-DEC-1998 (first entry)

Hormone-regulated Repro-PC-1.0 polypeptide.

Repro-PC-1.0; prostate cancer; LNCaP; hormone-regulated gene;

KW human; synaptotagmin.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Region 150..252  
 FT /note="repeat region"  
 FT Region 276..396  
 FT /note="repeat region"  
 FT Domain 15..37  
 FT /note="putative transmembrane domain"

W09839661-A1.

11-SEP-1998.

06-MAR-1998; 98NC-US04519.

15-MAY-1997; 97US-0047811.

07-MAR-1997; 97US-0041246.

(REPR-) REPROGEN INC.

French CK, Yamamoto KK;

WPI: 1998-506379/43.

N-PSDB: AAV57327.

Identification of hormone-regulated traits, e.g. in cancers - by exposing grafts of biological material to different hormonal environments in animals of different reproductive states

Example 1; Page 59-60; 85pp; English.

This polypeptide is encoded by a newly isolated hormone-regulated gene, termed Repro-PC-1.0 (see AAV57327), isolated from prostate cancer cells. The polypeptide (calculated mol.wt.48,070, pI 8.83) contains 2 copies of a 116-amino acid repeat that show homology to the C2 regulatory domain of calcium-dependent isoforms of protein kinase C, and to isoforms of synaptotagmin. Repro-PC1.0 expression is specifically and differentially up-regulated in LNCaP tumour cells. The invention provides methods of identifying hormone-regulated traits in a cell. The methods involve culturing the cell as a graft in 2 different hormonal environments and determining whether expression of the trait differs in the 2 grafts. The methods can be used to identify hormonally-regulated traits and hormonally-regulated genes for use as targets for therapeutic intervention in disease states, particularly cancers.

Sequence 425 AA:

Query Match 100.0%; Score 2201; DB 19; Length 425;

Best Local Similarity 100.0%; Pred. No. 2, 8e-196;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPITTSREEDPEIPVYVIGFSAFGLVTVSLFAMICCORSSKSNKTPPYKRVHVLKGV 60  
 DB 1 mapitstreedeipvtvgifsafiglvfsvlfawiccgqksksnktppykrtvhlkgy 60  
 QY 61 DIYPENLNKKKFGADDNENKRPAPVKNLSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120  
 DB 61 diypenlnskkfgaddnknknpavpknslhldlekrdngnfpktnlkpspsdlen 120  
 QY 121 ATPKFLGEGKESVPSLSKSTSLTSEEKOEKLTLPFSLFENFEKRAFYVNINKEARGL 180  
 DB 121 atpkflgegesvpslskstsltseekoeikltlpfslfeynfekafvnlkeargl 180  
 QY 181 PAMDEOSMTSDPYIKMTILPEKKHKVKTIVLAKTIDPAFDETFYFGIPYQIOELAHF 240  
 DB 181 pamdeqsmtsdpyikmtllpekhhkvktvrlktdpafdefctfygipqyqelahlf 240

OY	241	TLISDRFSRDDITIGEVILPLISGIELSEBKMILMNRILIRNNKSSGREGHLLISLCYOST	300
OY	241	TLISDRFSRDDITIGEVILPLISGIELSEBKMILMNRILIRNNKSSGREGHLLISLCYOST	300
Db	241	TLISDRFSRDDITIGEVILPLISGIELSEBKMILMNRILIRNNKSSGREGHLLISLCYOST	300
OY	301	TNTLIVVYLKARHLEPKSDVSGSLDPYVKVKNLYHAKKRISKSKTKHYKCPNVAENLEFVF	360
Db	301	TNTLIVVYLKARHLEPKSDVSGSLDPYVKVKNLYHAKKRISKSKTKHYKCPNVAENLEFVF	360
OY	361	DIPCEGLIEDISVEFLVLDSERGSRNEVIGQLVLGAAGSTGGEHMKELICDYPRRQIAKWH	420
Db	361	dipcegliedisveflvldeersgrnevigqlvlgaagstggehmkelidcyprrqiaakh	420
OY	421	VLCDG 425	
Db	421	VLCDG 425	
RESULT	3		
AAM39577	ID	AAM39577 standard; Protein; 431 AA.	
XX	AC	AAM39577;	
XX	DT	22-OCT-2001 (first entry)	
XX	DE	Human polypeptide SEQ ID NO 2722.	
XX	KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
XX	KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
XX	KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
XX	KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
XX	KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;	
XX	OS	leukaemia.	
XX	OS	Homo sapiens.	
XX	PN	WO200153312-A1.	
PD	26	JUL-2001.	
XX	26	DEC-2000; 2000WO-US34263.	
XX	21	JAN-2000; 2000US-0486725.	
XX	25	APR-2000; 2000US-0552317.	
PR	09	JUL-2000; 2000US-0596042.	
PR	19	JUL-2000; 2000US-0620312.	
PR	03	AUG-2000; 2000US-0653450.	
PR	14	SEP-2000; 2000US-0662191.	
PR	19	OCT-2000; 2000US-0693036.	
PR	29	NOV-2000; 2000US-0727344.	
PA	(HYSE-)	HYSEQ INC.	
PI	Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;		
XX	WPI: 2001-442253/47.		
DR	N-PSDB: AAI58733.		
XX	Novel nucleic acids and polypeptides, useful for treating disorders		
XX	such as central nervous system injuries -		
PS	Example 4; SEQ ID NO 2722; 10078bp; English.		
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and		
CC	the encoded polypeptides (AAM3642-AAM42213) with nootropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localized neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		

[illegible]





KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 PN MO200058473-A2.  
 PD 05-OCT-2000.  
 XX  
 PE 31-MAR-2000; 2000MO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 PI Shinkets RA, Leach M;  
 DR WPI: 2000-602362/57.  
 DR N-PSDB; AAC76182.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 11: Page 2647-2648; 5507pp; English.  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;  
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;  
 CC osteoplastic; anticonvulsant; antiallergic; immunosuppressive;  
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatologic; immunosuppressive;  
 CC antihistaminic; antibacterial; antiviral; antitumor; antirheumatic;  
 CC antithyroid; and antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 CC  
 XX  
 SQ Sequence 474 AA:

Query Match 21.1%; Score 464.5; DB 21; Length 474;  
 Best Local Similarity 30.3%; Pred. NO. 1.3e-34;  
 Matches 133; Conservative 71; Mismatches 152; Indels 83; Gaps 15;

QY 35 WICGQRKSSKSNKTPYKVFHVLKGVDPENLNSKKRFGADCKEYKKNPAVKKNSLHL 94  
 DB 60 WMASSSSddqgss-----vhaasevpltpfnspddtrissds-----kslysl 105  
 QY 95 -----DLKKRLNGNPKTNLKGSPDLENATPKLF--LEGEKESVSPESLKS----- 142  
 DB 106 Lrrisslesr-----fppsp--ltdikpiefgvisakkepqpsvlttrtynpd 152

QY 143 -----TSLTSEE--KOEKLGTEFFSLENYFEKRAVNIKEARCLP 181  
 DB 153 yfkfkephlysidssddvdsldteellskysqgmhfstqydlhnhltvviaeardp 212  
 QY 182 AM-----DEOSMT-SDPYTKMTILPEKKHKVKTFRVLRKTLDPAFDETFYVGIPTQIOE 235  
 DB 213 pphshdgsqdmahsnpyvkicllpddknskqykvkrkckqpfceerylf-eipfleaqr 271  
 QY 236 LAHFTILSPDEFSRDDITGEVLIPLSGIELSEG---KMLMRELIKNNVKKSSGREL 291  
 DB 272 rclltlvdfdkfsfhcykvsvplceavdlvqgymwkal-----psqneveigel 325  
 QY 292 LISLQSTNTNLTIVYVLRKARLRKSDVSGISDPYKVNLYAKKRISKKTIVKCTPN 351  
 DB 326 llsinypsagrlnvavlrakqllqtdvsqsdpfvrlqdlvnglkivaktktsflrgld 385  
 QY 352 AVFNELEVFEDIPCEGLIEDISVEPLVLDSEKRSNEYIGOLVGAAGG-TGGEHKKETCD 410  
 DB 386 pfynesfsikypqeelenaslvfvgfhmkssndfgrivlqyvsqpselnhrmln 445  
 QY 411 YPRRQIAKWHVL-----CD 424  
 DB 446 thrlaveqwhsirsraecd 464  
 XX  
 RESULT 8  
 AAM40253  
 ID AAM40253 standard; Protein: 474 AA.  
 XX  
 AC AAM40253;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 3398.  
 XX  
 KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW autoimmune lateral sclerosis; Shy-Drager Syndrome; Chemotactic;  
 KW leukoemetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW  
 XX  
 OS Homo sapiens.  
 PN MO200153312-A1.  
 PD 26-JUL-2001.  
 XX  
 PE 26-DEC-2000; 2000MO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR N-PSDB; AAI59409.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 5: SEQ ID NO 3398; 10078bp; English.  
 XX

[illegible]

```

Db 305 rllltvvdcktsrhvvgvplcevdllvkgghwkkali-----pssqnevevelgel 358
OY 292 LISLCYOSTNTTLTVVYLKARHLPKSDVSGLSDPYVKVNLTHAKKRISKKTTHVKKCTN 351
Db 359 llslnypsgirlvndqllqldvsgsqdptvklqvlhglklykktksfligld 418
OY 352 AVFNELEVPDPCBGEEDISVEFLVDSESGSRNEVIGOLVGAAGC-TGGEHKEICD 410
Db 419 pfynestsfkvpgeelensalvltvfglnmksndfgrivlgyssspspephwrrmln 478
OY 411 YPRROIAKMHVL-----CD 424
Db 479 thraveqwhslrsaeed 497

```

## RESULT 10

AAW25032 standard; Protein; 412 AA.

AAW25032:

15-OCT-1997 (first entry)

Human Doc2-beta colon cancer protein.

Doc2-beta; Colon cancer; C2 region; neoplasia; tumour; recombinant;

diagnosis; human cell line SW480; ss.

Homo sapiens.

JP09154586-A.

17-JUN-1997.

08-DEC-1995; 95JP-0320881.

08-DEC-1995; 95JP-0320881.

(SHIO ) SHIONOGI & CO LTD.

WPI: 1997-367066/34.

N-PSDB: AAT79627.

New human colon cancer protein Doc2 beta - has two C2 regions

Claim 1; Page 6-8; 10pp; Japanese.

AAW25032 is the Doc2-beta protein containing two C2 regions, it was isolated from the human colon cancer cell line SW480. The protein and cDNA encoding it are useful for the production of products for the diagnosis of colon cancer. A vector and transformed host cell are also claimed and the protein can be produced recombinantly.

Sequence 412 AA:

Query Match 19.4%; Score 428; DB 18; Length 412;  
Best Local Similarity 33.3%; Pred. No. 2.0e-31;  
Matches 118; Conservative 57; Mismatches 151; Indels 28; Gaps 8;

```

OY 84 KPAV-----PNSLIHLDLEKRDINGNFPPTNLKPG-SPSDLENATPKLLEGEKESVS 135
Db 61 rpvaagagrrpsdgaredde--dvddqllfgaygsppspqpsparppakpedepdadg 118
OY 136 PESLKSSTLSLSEKQEKLTGTFELEYENFERKAFVNIKEARGLPAMDROSTSPYTK 195
Db 119 yes-----ddctatgatlfdlslyldegnaahctfckakgkxpmhngl-addyk 167
OY 196 MTLIP--EKKHKVKTIVLKRITLDPAFDETFYFGIPYTOIOLALHTTILSPFRSRODI 253
Db 168 lhlppgasknkrktlrlntlnpwmcllyygltdedmlrktltsvdedckflhnef 227

```

```

OY 254 IGEVLIPLSGIELSRGK---MLMNR--IIRKRVKSSGRGELLISLCYSTNTTLTVV 308
Db 228 lgecvrplkklpnhctkfnctlekqlpvcktedksleergtllslkysqkglvyl 287
OY 309 LKARHLPKSDVSGLSDPYVKVNLTHAKKRISKKTTHVKKCTPAVAFNELEVPDPCBGE 368
Db 288 vrcablaamdangysdpyvktlylrpdvdkkskhktavkkklmpetneefcyeclhgdla 347
OY 369 DISVEFLVDSESGSRNEVIGOLVGAAGCTGGEHKEICDYPRROIAKMHVL 422
Db 348 kkslevtwdydglksndfignvvlghakgerlkhwfdclknkdrlertwll 401

```

## RESULT 11

AAR57421 standard; Protein; 704 AA.

AAR57421:

14-MAR-1995 (first entry)

Rabphilin-3A.

Low molecular weight; G protein; target protein; rab3a p25;

Rabphilin-3A; brain; nerve transmitter.

Homo sapiens.

JP06184199-A.

05-JUL-1994.

24-DEC-1992; 92JP-0344055.

24-DEC-1992; 92JP-0344055.

(EISA ) EISAI CO LTD.

WPI: 1994-252836/31.

N-PSDB: AAO67062.

Target protein of a low molecular G protein rabphilin-3A (RAB3A)

- found in the brain and involved in release of nerve transmitter

substance

Claim 1; Page 6-9; 9pp; Japanese.

This sequence represents a low molecular weight G protein target

protein, designated rab3a p25. Rabphilin-3A (Rab3A) is distributed

specifically in brain tissue and participates in the release of

nerve transmitter substance and is useful in the study of its

secretion.

Sequence 704 AA:

Query Match 18.5%; Score 408; DB 15; Length 704;  
Best Local Similarity 29.4%; Pred. No. 4.2e-29;  
Matches 99; Conservative 65; Mismatches 145; Indels 28; Gaps 4;

```

OY 112 PGSPSDLENATPKLFLGGEKESVSPESLKSSTLSLSEKQEKLTGTFELEYENFERKAFV 171
Db 361 pgsyqaaaaapqpvvasarqppeedeanaasydsdeatlgaletfslylqdgnslh 420
OY 172 VNIKARGLPAMDROSTSPYIKMTLP--EKKHKVKTIVLKRITLDPAFDETFYFGIP 229
Db 421 clikakgkxpmhngl-addykklhllpgasknkrktlrlntlnpwmcllyhngit 479
OY 230 YVQIOELALHFTTILSPFRSRODIIGEVLIPLSGTLESP----- 268
Db 480 dedmkrklrlsvdedckfghnetfgetfslkklpqrkhnfnclervlpmkragtq 539
OY 269 ---GKMLMNRILIRKRVKSSGRGELLISLCYSTNTTLTVVYLKARHLPKSDVSGLSDP 325

```



```

Db      540 sargmalveeegver-igdieergkllvalmystgtg9glivgllircvhlhaamdangysdp 598
QY      326 YKRVNLYHAKKRISKKTTHVKKCTPNVAVNNELEFVDPICEGLEDISVEFLVLDSEGSRN 385
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      599 fvklylkipdmgkakkhltqkkkltlnpeneefdydikhslakksldisvwdydlgksn 658
QY      386 EVIGQLVLGAAGTGTGEHMKETCDYPRROIARHVL 422
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      659 dyigcgqiglsakgerlkhwyec[knkdckierwhql 695

```

## RESULT 12

```

AAV27275
ID      AAV27275 standard; Protein; 405 AA.
XX
AC      AAV27275;
XX
DT      22-OCT-1999 (first entry)
XX
DE      Mouse Doc2alpha protein sequence.
XX
KM      Transgenic mouse; Doc2alpha; nervous disease; endocrine system.
XX
OS      Mus sp.
XX
PN      JP1196709-A.
XX
PD      27-JUL-1999.
XX
PF      07-JAN-1998; 98JP-0001925.
XX
PR      07-JAN-1998; 98JP-0001925.
XX
PA      (SHIO ) SHIONOGI & CO LTD.
XX
DR      WPI; 1999-496580/42.
XX
DR      N-PSDB; AA200321.
XX
PT      New transgenic mouse in which Doc2alpha is deleted - used for
PT      diagnosis of nervous system and endocrine disorders
XX
PS      Disclosure; Page 8-10; 13pp; Japanese.
XX
CC      The invention provides a transgenic mouse in which the sequence encoding
CC      the Doc2alpha protein is deleted. The mouse is useful as an experimental
CC      animal for the elucidation of causes of diseases in nervous and endocrine
CC      systems and the development of treating methods. The present sequence
CC      represents the mouse Doc2alpha protein sequence.
XX
SQ      Sequence 405 AA:

```

```

Query Match      17.9%; Score 393.5; DB 20; Length 405;
Best Local Similarity 31.2%; Pred. No. 4.1e-28;
Matches 106; Conservative 55; Mismatches 144; Indels 35; Gaps 7;

```

```

QY      112 PPSPSULENATKRLPLEGKESVSPESLKSSTLSNSEEOKLGTLPFSLFNFEKRAV 171
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      58 pahlaplalappaal1----galtcpddgaevdsydd--ltalgltelfllydgaasmh 112
QY      172 VNKEARGLPAMDEQSMSTDPYIKMTLPE--KKHKVTRVLRKTLDPAFDEFTFYGIP 229
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      113 clllrakgkmpdngfi-adpyvkhilpysackanklktqntlnpwnneelysgilt 171
QY      230 YVQIQELALHFTILSPDRSRDDIIGEVLIPLSGIELSEGG--MLMNRRI----- 277
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      172 dddilthkvlrvsvcdedkshnefiegeltvplrrllkpsqkhhnclerqvpfpssms 231
QY      278 -----IKRNRKSSG-----RGELLISLCYOSTTNTLTFLVVYLKARHLPKSDVSG 322
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      232 aalrtgscylkeleagaegpgjllleergpillsysstrhgllvgivrcahlhaamdvngy 291

```

```

QY      323 SDPYKVNLYHAKKRISKKTTHVKKCTPNVAVNNELEFVDPICEGLEDISVEFLVLDSENG 382
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      292 sdpykvtlylrpdvdkskkktcvtckkltlnpeneefyielstlaktletvtwdydlg 351
QY      383 SRNEVIGQLVLGAAGTGTGEHMKETCDYPRROIARHVL 422
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      352 kandfigvslgpargeagqkhwndolhgpdtalerwhcl 391

```

## RESULT 13

```

AAW01114
ID      AAW01114 standard; Protein; 400 AA.
XX
AC      AAW01114;
XX
DT      26-NOV-1996 (first entry)
XX
DE      Brain-specific protein, Doc2.
XX
KM      Doc2; brain-specific; neurotransmitter; diagnostic; therapeutic;
XX      cerebral nervous system.
XX
OS      Homo sapiens.
XX
PN      JP08168385-A.
XX
PD      02-JUL-1996.
XX
PF      22-DEC-1994; 94JP-0320775.
XX
PR      21-OCT-1994; 94JP-0257205.
XX
PA      (SHIO ) SHIONOGI & CO LTD.
XX
DR      WPI; 1996-357242/36.
XX
DR      N-PSDB; AAT40760.
XX
PT      DNA encoding brain-specific protein, Doc2 - is useful in elucidation
PT      of neuro:transmitter releasing mechanisms
XX
PS      Claim 1; Page 10-12; 15pp; Japanese.
XX
CC      The present sequence is that of a brain-specific protein, Doc2. It is
CC      useful in the elucidation of neurotransmitter releasing mechanisms. The
CC      gene (AAT40760) and anti-Doc2 antibodies are useful for the development
CC      of diagnostic and therapeutics for the cerebral nervous system.
XX
SQ      Sequence 400 AA:

```

```

Query Match      17.5%; Score 384.5; DB 17; Length 400;
Best Local Similarity 32.0%; Pred. No. 2.8e-27;
Matches 102; Conservative 50; Mismatches 136; Indels 31; Gaps 6;

```

```

QY      133 SVSPESLKSSTLSNSEEOKLGTLPFSLFNFEKRAVYVNKEARGLPAMDEQSMSTSP 192
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      70 atcpedgaevdsydd--atalgltelfllydraactlhwcllrakgkmpdngfi-adp 127
QY      193 YIKMTLPE--KKHKVTRVLRKTLDPAFDEFTFYGIPYVQIQELALHFTILSPDRFSR 250
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      128 ykklhllpysackanklktqntlnpwnneelysgiltdddilthkvlrvsvcdedksh 187
QY      251 DDILGVLPLPSGIELSEGG--MLMNRRI-----IKRNRKSSG-- 287
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      188 netfigeltvplrrllkpsqkhhnclerqvpilaspsmsaalrgiscylkdleagaegqg 247
QY      288 ----RGELLISLCYOSTTNTLTFLVVYLKARHLPKSDVSGSDPYKVNLYHAKKRISKKT 343
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      248 lilegrllllslsystrrgllvgllrcahlhaamdvngysdpyvkhilpdpvdkkshkt 307
QY      344 HVAKCTPNVAVNNELEFVDPICEGLEDISVEFLVLDSEGRSNEVIGQLVLGAAGTGE 403
      :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
Db      308 cvkkcllnefneeffyielstlaktletvtwdydlgskndffigvslgpargeark 367

```



CC sequences. The nucleic acids encoding the signal peptide can be used for  
 CC directing extracellular secretion of a polypeptide or the insertion of a  
 CC polypeptide into a membrane, or importing a polypeptide into a cell.  
 XX  
 SQ Sequence 72 AA;

Query Match 16.9%; Score 373; DB 20; Length 72;  
 Best Local Similarity 98.6%; Pred No. 2.5e-27;  
 Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPITTSREEFDEIPYVGIFSAFGLVFTVSLFAWICQQRKSSKSNKTPPYKRFVHLKGV 60  
 |||  
 Db 1 mapittsreefdeipvtvgifsafglvftvslfawicqqrksksknktppykrfvhlxgv 60

QY 61 DIYPENLNSKKK 72  
 |||  
 Db 61 diypenlnskkk 72

Search completed: December 12, 2001, 10:54:23  
 Job time: 34 sec

•  
•  
•  
•



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2001, 10:53:49 ; Search time 16 Seconds  
(without alignments)  
2023.387 Million cell updates/sec

Title: US-09-680-121-2

Perfect score: 2201

Sequence: 1 MAP1TTSREEFDEIPYVGI.....KEICDYPRIQIAKWHVLCDS 425

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query length	ID	Description
1	2006	91.1	425	2	IS9355
2	627.5	28.5	403	2	S58400
3	623.5	28.3	474	1	BMFEYS
4	621	28.2	403	2	S3318
5	603.5	27.4	422	1	BMRT2Y
6	603.5	27.4	422	2	A55417
7	603	27.4	511	2	S58399
8	594.5	27.0	537	2	JH0415
9	594	26.9	441	2	A40707
10	592	26.9	498	2	PC6300
11	581	26.4	422	1	BMHUIY
12	581	26.4	422	2	AA5486
13	579.5	26.3	427	2	JH0413
14	577	26.2	421	2	S09595
15	575	26.1	424	2	I51210
16	574.5	26.1	439	2	JH0414
17	573.5	26.1	386	2	I59387
18	547	24.9	279	2	S58402
19	544.5	24.7	390	2	T28967
20	520	23.6	588	2	A53563
21	474	21.5	474	2	S68695
22	430.5	19.6	412	2	JC4921
23	415	18.9	257	2	T16355
24	408	18.5	704	2	AA8097
25	394.5	17.9	681	2	JX0338
26	393.5	17.9	684	2	I58166
27	390.5	17.7	315	2	T32059
28	384.5	17.5	400	2	JC2473
29	377.5	17.2	504	2	T33485

30	361.5	16.4	355	2	S58401	synaptotagmin VIII
31	332	15.1	1021	2	S44644	hypothetical prote
32	323	14.7	387	2	JC7398	double C2 protei
33	217.5	9.9	448	2	T24770	hypothetical prote
34	184.5	8.4	743	2	T00634	hypothetical prote
35	179.5	8.2	682	1	KIRBGC	protein kinase C (
36	179.5	8.2	697	1	KIRBGC	protein kinase C (
37	178.5	8.1	697	2	D24664	protein kinase C (
38	178.5	8.1	1199	2	T33337	hypothetical prote
39	177.5	8.1	697	1	KIRTCG	protein kinase C (
40	177.5	8.1	697	2	JN0548	protein kinase C (
41	174.5	7.9	1212	2	T00332	hypothetical prote
42	174	7.9	861	2	T15903	protein kinase C h
43	165.5	7.5	137	2	S58403	synaptotagmin VIII
44	164	7.5	672	1	KIRBC	protein kinase C (
45	163.5	7.4	829	2	S58888	ins P4-binding pro

## ALIGNMENTS

```

RESULT 1
159355
synaptotagmin IV - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #ext_change 13-Aug-1999
C:Accession: I59355; I58163
R:Vicitan, L.; Lim, I.K.; Ferguson, G.; Tocco, G.; Baudry, M.; Herschman, H.R.
Proc. Natl. Acad. Sci. U.S.A. 92, 2164-2168, 1995
A:Title: Synaptotagmin IV is an immediate early gene induced by depolarization in PC1
A:Reference number: I59355; MUID:95199312
A:Accession: I59355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-425 <RES>
A:Cross-references: GB:I38247; NID:9598376; PIDN:AAA67327.1; PID:9598377
R:Ollrich, B.; Li, C.; Zhang, J.Z.; McMahon, H.; Anderson, R.G.; Geppert, M.; Sudhof,
Neuron 13, 1281-1291, 1994
A:Title: Functional properties of multiple synaptotagmins in brain.
A:Reference number: I58163; MUID:95085772
A:Accession: I58163
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-425 <RES>
A:Cross-references: EMBL:U14398; NID:9550453; PIDN:AAA6519.1; PID:9550454
C:Superfamily: synaptotagmin; protein kinase C C2 region homology
F:147-262/Domain: protein kinase C C2 region homology <KC2A>
F:281-396/Domain: protein kinase C C2 region homology <KC2B>

Query Match          91.1%; Score 2006; DB 2; Length 425;
Best local Similarity 89.9%; Pred. No. 1.5e-131;
Matches 382; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

QY 1 MAP1TTSREEFDEIPYVGIFFSAFGLFTVSLPAWICCKSSKSNTPYKRVHVLKGV 60
DB 1 MAP1TTSREEFDEIPYVGIFFSAFGLFTVSLPAWICCKSSKSNTPYKRVHVLKGV 60
QY 61 DIPEPMLNSKKKGGADKNEKKNKPAVPKNSLHLDLEKRDLDNPFKTNLPGSPDLEN 120
DB 61 DIPEPMLNSKKKGGADKNEKKNKPAVPKNSLHLDLEKRDLDNPFKTNLPGSPDLEN 120
QY 121 ATPKFLLEGKESVSPESLKSTSLTSEKOEKLGTLFSLSEYNFERKAFVNIKEARGL 180
DB 121 VPKFLPETEKEAVSPESLKSTSLTSEKOEKLGTLFSLSEYNFERKAFVNIKEARGL 180
QY 181 PAMDEQSMSTDPYIKMTILPEKKHKVTRVLRKTLDPAPPEFTFGIPYTOIELALHF 240
DB 181 PAMDEQSMSTDPYIKMTILPEKKHKVTRVLRKTLDPAPPEFTFGIPYTOIELALHF 240
QY 241 TILSPFRSHDDILGELVLPISGIELSEKMLMNRILKRNVRKSSGREGLLISLCYQST 300
DB 241 TILSPFRSHDDILGELVLPISGIELSEKMLMNRILKRNVRKSSGREGLLISLCYQST 300

```

QY 301 TMTLVVVLKARHLKPKSDVSGLSDEPVVKNLYHAKKRISKKTTHVKKCTPNVAVNELPVF 360  
 Db 301 TMTLVVVLKARHLKPKSDVSGLSDEPVVKNLYHAKKRISKKTTHVKKCTPNVAVNELPVF 360  
 QY 361 DIPCEGLEDISVEFLVLDSEKSRNEVIGOLVLAAGAEKGEGHMKETICDPYPRROIAKM 420  
 Db 361 DIPCEGLEDISVEFLVLDSEKSRNEVIGOLVLAAGAEKGEGHMKETICDPYPRROIAKM 420  
 QY 421 VTCGG 425  
 Db 421 MLCDG 425

## RESULT 2

S58400  
 Cellulagin II sytVII - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 13-Aug-1999  
 C:Accession: S58400  
 R:Li, C.; Ullrich, B.; Zhang, J.Z.; Anderson, R.G.W.; Brose, N.; Suedhof, T.C.  
 Nature 375, 594-599, 1995  
 A:Title: Ca(2+)-dependent and -independent activities of neural and non-neural synaptotagmin  
 A:Reference number: S58399; MUID:95312080  
 A:Accession: S58400  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-403 <LIC>  
 A:Cross-references: EMBL:U20106; NID:9643655; PIDN:AAA87725.1; PID:9643656  
 C:Superfamily: synaptotagmin; protein kinase C C2 region homology  
 F:129-243/Domain: protein kinase C C2 region homology <KC2A>  
 F:260-375/Domain: protein kinase C C2 region homology <KC2>

## Query Match

Best Local Similarity 28.5%; Score 627.5; DB 2; Length 403;  
 Matches 153; Conservative 72; Mismatches 127; Indels 71; Gaps 11;

QY 27 VETVSL-----FAMIC-C-CORRKSNNKTPPYKFFVHLKGVLDIPEMLNKKKFGAD-- 77  
 Db 23 IIVTSLVTVTLGCLHMCORLGRKRYK-----NSLETVGTPDDSG 62  
 QY 78 -----KNEVK-----NKPAVPKNSLHLDLEKRDNGNFPKTNLKPSPDLNATPK 124  
 Db 63 RGRGEKKAKILPAGKAVNTAPVPGOTPHDESDR---TEPRS-----SVSDLV---- 109  
 QY 125 LFLGEEKSVSPSLKSTSLTSEE-----KOEKLTFLFLEYNFERKAFVYNIKEARG 179  
 Db 110 -----SLTSEMLMLSPGSEDEAHGCSRENIGRQFSVGNFQESTLTVKVKAAE 161  
 QY 180 LPAWDEQSMSTDPYIKMTILPEKKHKVTRVLRKTLDPAPDPTFFGIPYQIOCLAH 239  
 Db 162 LPAKD-FSGSDSPFVYIYLLPDKKKHLETKYKRNLMNHNTEFLFEGFYEYVQRIIX 220  
 QY 240 FTLLSDRESRDITIGEVILPISGLESEGKMLNREIRKNNVKKSGRELLSLCTQS 299  
 Db 221 LQVLDHRSRNPJGEVSIPLAKVDLTO--MOTFWMDLPCSDSGSRELLSLCTQNP 278  
 QY 300 TMTLVVVLKARHLKPKSDVSGLSDEPVVKNLYHAKKRISKKTTHVKKCTPNVAVNELPV 359  
 Db 279 SANSITVNIKAKNLKAMDIGTSPYKVMYKDKRVEKKKTVYKRNLMNIFNESPA 338  
 QY 360 FDIPEGLEDISVEFLVLDSEKSRNEVIGOLVLAAGAEKGEGHMKETICDPYPRROIAKM 419  
 Db 339 FDIPTERKRETTIITVMDKDLISRDVIGIKIYLSMKSQGEVKKHMDIARQPVQW 398  
 QY 420 HVL 422  
 Db 399 HOL 401

RESULT 3  
 BMFFSY

synaptotagmin - fruit fly (Drosophila melanogaster)

N:Alternate names: p65  
 C:Species: Drosophila melanogaster  
 C>Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 22-Jun-1999  
 C:Accession: B39052  
 R:Bern, M.S.; Johnston, P.A.; Ozeccelik, T.; Jahn, R.; Francke, U.; Suedhof, T.C.  
 J. Biol. Chem. 266, 615-622, 1991  
 A:Title: Structural and functional conservation of synaptotagmin (p65) in Drosophila  
 A:Reference number: A39052; MUID:91093190  
 A:Accession: B39052  
 A:Molecule type: mRNA  
 A:Residues: 1-474 <PER>  
 A:Cross-references: GB:M55048; GB:J05711; NID:9158526; PIDN:AAA28925.1; PID:9158527  
 C:Comment: Synaptotagmins are a major component of synaptic vesicle membranes and are  
 C:genetics:  
 A:Gene: Flybase:syt  
 A:Cross-references: Flybase:FBgn0004242  
 A:Map position: 23b  
 C:Superfamily: synaptotagmin; protein kinase C C2 region homology  
 C:Keywords: duplication; membrane trafficking; phospholipid binding; synaptic vesicle  
 F:1-107/Domain: Intravesicular #status predicted <INT>  
 F:108-134/Domain: Transmembrane #status predicted <TM>  
 F:135-474/Domain: extravesicular #status predicted <EXT>  
 F:186-434/Region: phospholipid binding #status experimental  
 F:186-300/Domain: protein kinase C C2 region homology <KC2A>  
 F:319-434/Domain: protein kinase C C2 region homology <KC2B>

## Query Match

Best Local Similarity 28.3%; Score 623.5; DB 1; Length 474;  
 Matches 152; Conservative 66; Mismatches 144; Indels 73; Gaps 9;

QY 3 PITTSREEDLEIPYV-----GIFSAFLVFTVSLFAMICORRKSNNKTPPYK 52  
 Db 88 PVIKRKEHGEVTEVIAERTGLPTMGVVAIILVFLVVGIIFFCVRRFLKRRTK--- 144  
 QY 53 FVHVLKGVLDIPEMLNKKKFGADCKNEKPAVPKNSLHLDLEKRDNGNFPKTNLKP 112  
 Db 145 -----DGKGRKG-----VDMSVQLGSAAYEKVQF 170  
 QY 113 GSPSDLENATPKLFLGEEKSVSPSLKSTSLTSEKOEKLTFLFLEYNFERKAFVY 172  
 Db 171 DMBELTEANE-----EGDE-----DKSQKILGRINFLKEVDFNSLAV 211  
 QY 173 NIKERAGLPAMDEQSMSTDPYIKMTILPEKKHKVTRVLRKTLDPAPDPTFFGIPYQ 232  
 Db 212 TVIQAEELPAID-MGCTSDPYKVVLLPDKKKHLETKYKRNLMNHNTEFLFEGFYEYVQ 270  
 QY 233 IOELAHFTLLSPRESRDITIGEVILPISGLESEGKMLNREIRKNNVKKSGRELLSLCT 289  
 Db 271 ANMKTVLRAIDPFDEFKSHDQIGEVKVPCLCTIDLAQ-TIEENRDL--SVBEGQDEKLG 327  
 QY 290 ELLISLCQSTNTITLVVVLKARHLKPKSDVSGLSDEPVVKNLYHAKKRISKKTTHVKKCT 349  
 Db 328 DICFSRTVPTAGKTVLLEAKNLKMDVGLSDPYKVIATIMQNKRRKKKTSVKKCT 387  
 QY 350 PNAVNELFVDIPCEGLEDISVEFLVLDSEKSRNEVIGOLVLAAGAEKGEGHMKETIC 409  
 Db 388 LNPYVNESFSEVPEQOKICLVVTVVDYDRIGTSPIGRICLQCMGTGELRWSDML 447  
 QY 410 DYPRROIAKMHVLC 424  
 Db 448 ASRPRIQWHTLKD 462

## RESULT 4

S33318  
 synaptotagmin - longfin squid  
 C:Species: Loligo pealeii (Longfin squid)  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 13-Aug-1999  
 C:Accession: S33318  
 R:Bommet, K.; Charlton, M.P.; DeBello, W.M.; Chin, G.J.; Betz, H.; Augustine, G.J.  
 Nature 363, 163-165, 1993



Fri Dec 14 10:32:17 2001

us-09-680-121-2.rpr

Page 4

```

Db      106 - KMANNMKMGCGDDDDA----- -FTGLEG- 130
OY      123 KRLLEGGKEVSPESIKSSTSLTSEKOEKLTTFEFSLEVFERRKAVVVIKEARGLA 182
Db      131 -----EGGE----- EKERENLNGKIOFSLDYVFOQMOITVGLDIAELPA 171
OY      183 MDEOSMTSDPYIKMTILPEKHKRVKTRVLKTLDPADFETTFYVGIPIYTOIOELAHFTI 242
Db      172 LD -MGSITDPYKVFLLPDKKKKEITKRVKTKLNPENETTFE -KVPOEIGGKTLVMAI 229
OY      243 LSPRESDPIDIGVLLPLSLGEL----SEKMLANREIKRNVKSSGREGELISLCYQ 298
Db      230 YDFRFKEKHIDIGVWKVPMYVLDGPIEBWDLQGE -KEEPR-- -LDITSLSRIV 284
OY      299 STTTLVLVVVLLAKNRIEKRSPVSGSLDPYKVVLLYAKKRIKSKTHYKCPNVAPELF 358
Db      285 PTAKRLTVCILTEAKNNLKKMDVGLSDPYVKHILMONGKRLLKMTKYKKTLNRYNESF 344
OY      359 VEDPDCGLEDISAEFLVLDSEKRSNEVITGQVLVGAAEGGGHHMEKIIDYRRRIAK 418
Db      345 SFELRFOIOIKOVVVVTLVDYDKLGNKNEAIGKIFVGSNATGTSLRHWSMDLANRPRIAQ 404
OY      419 WHVL 422
Db      405 WHSL 408

```

RESULT 7  
S58399 cellutagmin I svrVI - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Apr-1996 #sequence-revision 19-Apr-1996 #text-change 13-Aug-1999  
C:Accession: S58399  
R:Lil, C.; Ullrich, B.; Zhang, J.Z.; Anderson, R.G.W.; Brose, N.; Suedhof, T.C.  
Nature 375, 594-599, 1995  
A:Title: Ca(2+)-dependent and -independent activities of neural and non-neural synaptobrevin  
A:Accession: S58399; MUID:95312080  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-511 <LIC>  
A:Cross-references: EMLC:U20105; NID:9643653; PIDN:AA87724.1; PID:g643654  
C:Superfamily: synaptotagmin; protein kinase C C2 region homology <KC2A>  
F:224-337/Domain: protein kinase C C2 region homology <KC2A>  
F:356-471/Domain: protein kinase C C2 region homology <KC2B>

Query Match	27.4%	Score 603	DB 2	Length 511
Best Local Similarity	40.4%	Pred. No. 2,9e-34		
Matches 127	Conservative	64	Mismatches 103	Indels 20
			Gaps	6
QY	115	PSDLNATPKFLFEGEESVSPESLSKSSNLSLSEKQEKGTFFSELENEPRAAFVNI	114	
		!:::!		
Db	202	PSIRIGRIPELY---KQSYVDGEAKS-----EAKSGCGKINFLSRADYSESLIRI	251	
		!:::!		
QY	175	KARGLPAMDQSWSPSTPYKMTIIPEKKHKVTRVRLKTLDPAPETTPGIPYQIQ	234	
		!:::!		
Db	252	LKAFDLPAKDFCG-SSDPYVKIYLLPDKCKLQTRVIRKRLTNPEDENFHP-PPYIELA	309	
		!:::!		
QY	235	ELAHFTLISDRSDRIIGVLPISLGIELSGKRLMAREIKKNR-----KSSGRGE	290	
		!:::!		
Db	310	DKRLHLSVFEDRDRSRHDMIGEVIID-----NLFASDLSREISIMKDLYQVATSSSDVJGE	365	
		!:::!		
QY	291	LILSLCOSTNLTLYVYLKAHLRKRSDVSGSLSPYKYNLYIAKKRIKSKKTHVKKCP	350	
		!:::!		
Db	366	IMFSLCTIPFAGRLITLVKCNRLAAMDTITGSDPYKQSLCDDGRKLKRTTKTKNTL	425	
		!:::!		
QY	351	NAVENELFVDPIDCEGLIEDIISVEFLYLDSENGSNEYIGOLVUGAAACGTGGEHWEICD	410	
		!:::!		
Db	426	NPVYKEALIEDIPENNDDVSLIISVMIDYDRVGNHETIGCVGISAGELGRDHNMENLA	485	
		!:::!		
QY	411	YPRQIAKHHVYCD	424	

DB 486 YPKRPIAHWNCIAE 499

RESULT 8

JH0415

synaptotagmin o-p65-C - electric ray (Discozyge ommata)

N:Alternate names: synaptotagmin vesicle protein o-p65-C

C:Species: Discozyge ommata

C:Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 13-Aug-1999

C:Accession: JH0415

R:Wendland, B.; Miller, K.G.; Schilling, J.; Scheller, R.H.  
Neuron 6, 993-1007, 1991

A:Title: Differential expression of the p65 gene family.

A:Reference number: JH0413; MUID:91273991

A:Accession: JH0415

A:Molecule type: mRNA

A:Residues: 1-537 <MEN>

A:Cross-references: Gb:M64277; NID:9213113; PIDN:AAA9229.1; PID:9213113

A:Experimental source: electric organ

A:Subcellularly: synaptotagmin; protein kinase C C2 region homology

F:5578-68: glycoprotein, membrane protein; synaptic vesicle

F:230-343:omata: hydrophobic <HD>

F:362-477/Domain: protein kinase C C2 region homology <RC2A>

F:362-477/Domain: protein kinase C C2 region homology <RC2B>

Query Match	27.0%	Score 594.5	DB 2	Length 537
Best Local Similarity	31.5%	Pred. No. 1.2e-33		
Matches 156	Conservative 88	Mismatches 140	Indels 111	Gaps 19
QY	6	TSREPFDEIPVWVIFSAF-GLVET-VSLF-AMTCC-----	QKSSKSKNTPPY	51
Db	46	TRHIEDIVSLSLVIVTFGCIYLLGLVSLFYFWMKLCWIPMRDGLIPQRHDSOHN----	PH	102
QY	52	KTVH-----VLKGVDIYEPENLSSKKKGGADDKNEVKNRPAAVKNSLHIDLEKRDLN	102	
Db	103	QHIIHHHSHPTDLTVERVDCGP-----	MPERS-YLDLE	135
QY	103	GNEPXTNLK-----PGSPDL-----EN-----	A	121
Db	136	SFSPESGKLSQTSPPDIIVDTVSSGSKENNIPNAHSGQVSAAPRATRENSLRPRIPPOOLS		194
QY	122	TEKRLDEBEK-----SVSPESLKSSTSLTSEEGQEL-GLTFSLSEYNEKRAFY		171
Db	195	SEPTGQADKEVEYVTSIQIKPELTKQSRSDITPAKKHQVNCGRINELKATYTTDGLV		254
QY	172	VNIKARSLRPMDOASTSPDYIKYKTLTPREKKHAKVRYLRIKTPADREFTFTYTGIRT		231
Db	255	VKILKAALDRPAKQNGF-SDPYVYKLYLLRDPKKKKFOYKRYAKRLINPINEFQF-NYFN		312
QY	232	QYQELALHFTLSRDSRDLIDGVLPR--LSQTELSEKMMIMREITIKKNVSSGSG		285
Db	313	ELQNKRLFSFYSDREDSRHDLDGOVULDELLEPDSDEDTIYW-RDLLEATSKAD-LG		370
QY	290	ELLISLCYOSTTNTLVVVKARHNLPRKDVSGSLSDPVVKNLYNAKKRYSKKNHVKCF		349
Db	371	EINFSLCYLPAGLCTLTIIKATNKAKMLDGFSDPVPYKASLIDCBERRIKKTSIKKNT		430
QY	350	PAVAENELVFEDIPGSGEDIDVEFVLVDSSRGSRNVEIGOLVULAAEGGSGHNKEIC		409
Db	431	LNPNVYNLEVFEDIPENMEHNVNIIAAMDYDCLGHNEBVGIMCRVGNATDGGREHVNEML		490
QY	410	DYPRQALAKHIVLCLD	424	
Db	491	ANPKRPTEIOWHOLIE	505	

RESULT 9  
AA0707  
synapcotagmin - *Caenorhabditis elegans*  
N:Alternate names: ric-2  
C:Species: *Caenorhabditis elegans*  
C>Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 24-Sep-1999









Fri Dec 14 10:32:17 2001

us-09-680-121-2.rpt

Page 8



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2001, 10:54:49 ; Search time 13.23 Seconds

(without alignments)  
1177.819 Million cell updates/sec

Title: US-09-680-121-2

Perfect score: 2201  
Sequence: 1 MAPITRSREDFEIPYVGI.....KEICDYPROIQAKWHLCDG 425

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt.39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2006	91.1	425	1 SYT4_RAT	P50232 rattus norv
2	1988	90.3	425	1 SYT4_MOUSE	P40749 mus musculu
3	623.5	28.3	474	1 SYT5_DROME	P21521 drosophila
4	606.5	27.6	426	1 SYT5_APLCA	P41823 aplysia cal
5	604.5	27.5	422	1 SYT2_MOUSE	P46097 mus musculu
6	603.5	27.4	422	1 SYT2_RAT	P23101 rattus norv
7	594.5	27.0	537	1 SYT3_DISOM	P24507 discoppyge o
8	594	26.4	441	1 SYT1_MOUSE	P34693 caenorhabdi
9	581	26.4	422	1 SYT1_MOUSE	P46096 mus musculu
10	581	26.4	422	1 SYT1_BOVIN	P48018 bos taurus
11	580	26.4	422	1 SYT1_HUMAN	P21579 homo sapien
12	580	26.4	421	1 SYT1_RAT	P21707 rattus norv
13	579.5	26.3	427	1 SYT1_DISOM	P24505 discoppyge o
14	575	26.1	424	1 SYT1_CHICK	P47191 gallus gall
15	574.5	26.1	439	1 SYT62_DISOM	P24506 discoppyge o
16	573.5	26.1	386	1 SYT5_RAT	P47861 rattus norv
17	571.5	26.0	386	1 SYT5_HUMAN	O00445 homo sapien
18	520	23.6	588	1 SYT3_RAT	P40748 rattus norv
19	516	23.4	587	1 SYT3_MOUSE	O35681 mus musculu
20	408	18.5	704	1 RP3A_BOVIN	O06846 bos taurus
21	399.5	18.2	694	1 RP3A_HUMAN	O92210 homo sapien
22	393.5	17.9	684	1 RP3A_RAT	P47709 rattus norv
23	332	15.1	1021	1 RP17_CAEL	P41885 caenorhabdi
24	300.5	13.7	606	1 RP3A_MOUSE	P47708 mus musculu
25	179.5	8.2	682	1 KPCG_BOVIN	P05128 bos taurus
26	179.5	8.2	697	1 KPCG_RABIT	P10829 oryctolaqus
27	178.5	8.1	697	1 KPCG_HUMAN	P05129 homo sapien
28	177.5	8.1	697	1 KPCG_MOUSE	P05129 homo sapien
29	175	8.0	826	1 RSG5_HUMAN	O43374 homo sapien
30	173.5	7.9	834	1 RSG3_BOVIN	O28013 bos taurus
31	166	7.5	834	1 RSG3_HUMAN	O16644 homo sapien
32	164	7.5	672	1 KP4A_BOVIN	P04409 bos taurus
33	163	7.4	672	1 KP4A_RABIT	P10102 oryctolaqus

34	163	7.4	672	1 KP4A_RAT	P05696 rattus norv
35	163	7.4	834	1 RSG3_MOUSE	O60790 mus musculu
36	162	7.4	639	1 KPCL_DROME	P05130 drosophila
37	162	7.4	1634	1 KP3B_HUMAN	O00750 homo sapien
38	161	7.3	672	1 KP4A_HUMAN	P17252 homo sapien
39	159	7.2	672	1 KP4A_MOUSE	P20444 mus musculu
40	159	7.2	847	1 RSG2_RAT	O63713 rattus norv
41	158	7.2	671	1 KPCL_RABIT	P05772 oryctolaqus
42	158	7.2	673	1 KPCL_BOVIN	P05126 bos taurus
43	158	7.2	673	1 KPCL_RAT	P05773 oryctolaqus
44	153	7.0	671	1 KPCL_MOUSE	P04410 rattus norv
45	153	7.0	673	1 KPCL_MOUSE	P04411 mus musculu

## ALIGNMENTS

```

RESULT 1
SYT4_RAT
ID SYT4_RAT STANDARD: PRT: 425 AA.
AC P50232:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SYNAPTOTAGMIN IV (SYTIV).
GN SYT4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=95085772; PubMed=7993622;
RA Ullrich B., Li C., Zhang J.Z., McMahon H., Anderson R.G., Geppert M.,
RA Suedhof T.C.;
RT "Functional properties of multiple synaptotagmins in brain.";
RL Neuron 13:1281-1291(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95199312; PubMed=7892240;
RA Vician L., Lim I.K., Ferguson G., Tocco G., Baudry M., Herschman H.R.;
RT "Synaptotagmin IV is an immediate early gene induced by
RT depolarization in PC12 cells and in brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:2164-2168(1995).
CC -!- FUNCTION: MAY BE INVOLVED IN CA2+-DEPENDENT EXOCYTOSIS OF
CC SECRETORY VESICLES THROUGH CA2+ AND PHOSPHOLIPID BINDING TO THE C2
CC DOMAIN OR MAY SERVE AS CA2+ SENSORS IN THE PROCESS OF VESICULAR
CC TRAFFICKING AND EXOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SYNAPTIC
CC VESICLES.
CC -!- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14398; AA68519.1; -.
DR EMBL: U38247; AA67327.1; -.
DR HSSP: P21707; IRSY.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002149; IRI.
DR InterPro: IPR001565; Synaptotagmin.
DR Pfam: PF00168; C2_2.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGMIN.
DR SMART: SM00239; C2_2.
DR PROSITE: PS00499; C2_DOMAIN_1; 2.

```

DR PROSITE: PS50004; C2\_DOMAIN.2; 2.  
 KW Transmembrane; Repeat; Synapse.  
 FT DOMAIN 1 16 VESICULAR (POTENTIAL).  
 FT TRANSMEM 17 37 POTENTIAL.  
 FT DOMAIN 38 425 CYTOSOLMIC (POTENTIAL).  
 FT DOMAIN 147 297 PHOSPHOLIPID BINDING (PROBABLE).  
 FT DOMAIN 155 258 C2 DOMAIN 1.  
 FT DOMAIN 289 392 C2 DOMAIN 2.  
 SQ SEQUENCE 425 AA: 47685 MW: 6AC8BE0878936BB CRC64;

Query Match Best Local Similarity 91.1%; Score 2006; DB 1; Length 425;  
 Matches 38; Conservative 21; Mismatches 22; Indels 0; Gaps 0;

QY 1 MAPITTSREDEIPVYVIGFSAFGLVFTVSLFAMICCCORSSKSKNTPPYKFFVHLKGV 60  
 DB 1 MAPITTSREDEIPVYVIGFSAFGLVFTVSLFAMICCCORSSKSKNTPPYKFFVHLKGV 60  
 QY 61 DIYPENLSKKKFGADKNEKKNPAVKNSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120  
 DB 61 DIYPENLSKKKFGADKNEKKNPAVKNSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120  
 QY 121 ATPKLFLEGEKESVPELSKSTSLTSEKOEKLTGLFFSLEYNEFKKAFVNIKEARGL 180  
 DB 121 ATPKLFLEGEKESVPELSKSTSLTSEKOEKLTGLFFSLEYNEFKKAFVNIKEARGL 180  
 QY 121 VTPKLFETETKEVSPESLSKSTSLTSEKOEKLTGLFFSLEYNEFKKAFVNIKEARGL 180  
 DB 121 VTPKLFETETKEVSPESLSKSTSLTSEKOEKLTGLFFSLEYNEFKKAFVNIKEARGL 180  
 QY 181 PAMDOQMTSDPYIKMTLPEKKHKVTRVLRKTLDPADFETFFYGIPTQIOELALHF 240  
 DB 181 PAMDOQMTSDPYIKMTLPEKKHKVTRVLRKTLDPADFETFFYGIPTQIOELALHF 240  
 QY 241 TIISFDFESRDITIGVLIPLSGIESEGKMLNRETIKRNVRKSSRGELLISLCYST 300  
 DB 241 TIISFDFESRDITIGVLIPLSGIESEGKMLNRETIKRNVRKSSRGELLISLCYST 300  
 QY 241 TVLSFDFESRDITIGVLIPLSGIESEGKMLNRETIKRNVRKSSRGELLISLCYST 300  
 DB 241 TVLSFDFESRDITIGVLIPLSGIESEGKMLNRETIKRNVRKSSRGELLISLCYST 300  
 QY 301 TNTLTVVYLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCPNVAENLEFVF 360  
 DB 301 TNTLTVVYLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCPNVAENLEFVF 360  
 QY 361 DIPEGLEDISVEFLVLDSEGRSNEVIGOLVGAAGSTGEGHKKICDYPRIQIAKWH 420  
 DB 361 DIPEGLEDISVEFLVLDSEGRSNEVIGOLVGAAGSTGEGHKKICDYPRIQIAKWH 420  
 QY 421 VLCDG 425  
 DB 421 VLCDG 425

RESULT 2  
 SYTA\_MOUSE STANDARD: PRT: 425 AA.  
 ID SYTA\_MOUSE  
 AC P40749;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE SYNAPTOTAGMIN IV (SYTIV).  
 GN SYTA OR SYT3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCBL\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BAIB/C; TISSUE=Brain;  
 RA MEDLINE=9436712; PubMed=8058779;  
 RL Hilbush B.S., Morgan J.I.;  
 RL A third synapocotagmin gene, Syt3, in the mouse.\*;  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:8195-8199(1994).  
 CC -1- FUNCTION: MAY BE INVOLVED IN CA2+-DEPENDENT EXOCYTOSIS OF  
 CC SECRETORY VESICLES THROUGH CA2+ AND PHOSPHOLIPID BINDING TO THE C2  
 CC DOMAIN OR MAY SERVE AS CA2+ SENSORS IN THE PROCESS OF VESICULAR  
 CC TRAFFICKING AND EXOCYTOSIS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. SYNAPTIC

CC VESICLES.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY REGIONS OF THE NERVOUS  
 CC SYSTEM BUT IS UNDETECTABLE IN EXTRA NEURAL TISSUES.  
 CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb.slb.ch](mailto:license@isb.slb.ch)).  
 CC  
 DR EMBL: U10355; AAA20971.1; -;  
 DR HSSP: P21707; IRSY.  
 DR MGD: MGI:101759; Syt4.  
 DR InterPro: IPR000008; C2.  
 DR InterPro: IPR002149; LRI.  
 DR InterPro: IPR001565; SynaptoTagmin.  
 DR Pfam: PF00168; C2; 2.  
 DR PRINTS: PR00360; C2DOMAIN.  
 DR PRINTS: PR00399; SYNAPTOTAGMIN.  
 DR SMART: SM00239; C2; 2.  
 DR PROSITE: PS00499; C2\_DOMAIN.1; 2.  
 DR PROSITE: PS50004; C2\_DOMAIN.2; 2.  
 KW Transmembrane; Repeat; Synapse.  
 FT DOMAIN 1 16 VESICULAR (POTENTIAL).  
 FT TRANSMEM 17 37 POTENTIAL.  
 FT DOMAIN 38 425 CYTOSOLMIC (POTENTIAL).  
 FT DOMAIN 147 297 PHOSPHOLIPID BINDING (PROBABLE).  
 FT DOMAIN 155 258 C2 DOMAIN 1.  
 FT DOMAIN 289 392 C2 DOMAIN 2.  
 SQ SEQUENCE 425 AA: 47630 MW: D1900DA75C163821 CRC64;

Query Match Best Local Similarity 90.3%; Score 1988; DB 1; Length 425;  
 Matches 38; Conservative 21; Mismatches 24; Indels 0; Gaps 0;

QY 1 MAPITTSREDEIPVYVIGFSAFGLVFTVSLFAMICCCORSSKSKNTPPYKFFVHLKGV 60  
 DB 1 MAPITTSREDEIPVYVIGFSAFGLVFTVSLFAMICCCORSSKSKNTPPYKFFVHLKGV 60  
 QY 61 DIYPENLSKKKFGADKNEKKNPAVKNSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120  
 DB 61 DIYPENLSKKKFGADKNEKKNPAVKNSLHLDLEKRDNGNFPKTNLKPSPSDLEN 120  
 QY 121 ATPKLFLEGEKESVPELSKSTSLTSEKOEKLTGLFFSLEYNEFKKAFVNIKEARGL 180  
 DB 121 ATPKLFLEGEKESVPELSKSTSLTSEKOEKLTGLFFSLEYNEFKKAFVNIKEARGL 180  
 QY 121 VTPKLFETETKEVSPESLSKSTSLTSEKOEKLTGLFFSLEYNEFKKAFVNIKEARGL 180  
 DB 121 VTPKLFETETKEVSPESLSKSTSLTSEKOEKLTGLFFSLEYNEFKKAFVNIKEARGL 180  
 QY 181 PAMDOQMTSDPYIKMTLPEKKHKVTRVLRKTLDPADFETFFYGIPTQIOELALHF 240  
 DB 181 PAMDOQMTSDPYIKMTLPEKKHKVTRVLRKTLDPADFETFFYGIPTQIOELALHF 240  
 QY 241 TIISFDFESRDITIGVLIPLSGIESEGKMLNRETIKRNVRKSSRGELLISLCYST 300  
 DB 241 TIISFDFESRDITIGVLIPLSGIESEGKMLNRETIKRNVRKSSRGELLISLCYST 300  
 QY 241 TVLSFDFESRDITIGVLIPLSGIESEGKMLNRETIKRNVRKSSRGELLISLCYST 300  
 DB 241 TVLSFDFESRDITIGVLIPLSGIESEGKMLNRETIKRNVRKSSRGELLISLCYST 300  
 QY 301 TNTLTVVYLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCPNVAENLEFVF 360  
 DB 301 TNTLTVVYLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCPNVAENLEFVF 360  
 QY 361 DIPEGLEDISVEFLVLDSEGRSNEVIGOLVGAAGSTGEGHKKICDYPRIQIAKWH 420  
 DB 361 DIPEGLEDISVEFLVLDSEGRSNEVIGOLVGAAGSTGEGHKKICDYPRIQIAKWH 420  
 QY 421 VLCDG 425  
 DB 421 VLCDG 425

```

RESULT 3
SY65_DROME STANDARD: PRT; 474 AA.
AC P21521;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
GN SYNAPTOTAGMIN (P65).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91093190; PubMed=1840599;
RA Perin M.S., Johnston P.A., Oezcelik T., Jahn R., Francke U.,
RA Suedhof T.C.;
RT "Structural and functional conservation of synaptotagmin (p65) in
RT Drosophila and humans";
RL J. Biol. Chem. 266:615-622(1991).
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE.
CC -1- SUBUNIT: HOMODIMER OR HOMOTRIMER (POSSIBLE).
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES IN NEURONS.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, M55048; AAA28925.1; -.
CC PIR, B39052; BMFESY.
CC HSSP, P21707; IRSY.
CC FLYBASE: FBgn0004242; syt.
CC DR FLYBASE: P21707; IRSY.
CC DR InterPro: IPR000008; C2.
CC DR InterPro: IPR002149; LRI.
CC DR InterPro: IPR001565; Synaptotagmin.
CC DR Pfam: PF00168; C2; 2.
CC DR PRINTS: PR00360; C2DOMAIN.
CC DR PRINTS: PR00399; SYNAPTOTAGMIN.
CC DR SMART: SM00239; C2; 2.
CC DR PROSITE: PS00499; C2_DOMAIN_1; 2.
CC DR PROSITE: PS50004; C2_DOMAIN_2; 2.
CC KW Transmembrane; Repeat; Synapse.
CC FT DOMAIN 1 107 VESICULAR.
CC FT TRASNEM 108 134 POTENTIAL.
CC FT DOMAIN 135 474 CYTOPLASMIC.
CC FT DOMAIN 136 434 PHOSPHOLIPID BINDING (PROBABLE).
CC FT DOMAIN 206 296 C2 DOMAIN 1.
CC FT DOMAIN 339 430 C2 DOMAIN 2.
CC FT SEQUENCE 474 AA; 53278 MW; BF52A26EAF923F6F CRC64;

```

```

Query Match 28.3%; Score 623.5; DB 1; Length 474;
Best Local Similarity 34.9%; Pred. No. 1.5e-35;
Matches 152; Conservative 66; Mismatches 144; Indels 73; Gaps 9;

```

```

OY 3 PTTSEEDDEPTTV-----GIFSAGELVFTVSLFAMICQRRSSKSKTKPPYK 52
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 88 PYAKKEHGEVYVTEIARTGILPTGWVAAIILVLFVFGIIFCVRRRLKRRRK--- 144
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 53 FVHLKGVVDIPENLSKKRRFGADKREYKNKPAVPRKNSLHDLERKRDNGNFKTKLKP 112
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

DB 145 -----DGKSKGK-----YDMKSVQLLSGAYKEKVP 170
OY 113 GSRSDLENATPKFLGEGKESVPSLSKSTSLTSEKOEKIGTLFESLYNERRAFV 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 171 DMEELTENAE-----EGDE-----DKQSQKIGRLNFKLEYDNNSLAV 211
OY 173 NIKFARGLPAMDQSWTSPYIKMTLPEKKHKVKTARKTLDPAFDETFPPYGPYQ 232
   : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 212 TVIQAEELPALD-MGSTDSPYKAVILLPDKKKFELKVRKTLSPVFNFTFKSLPYAD 270
OY 233 IOELALHTLSFDRESRDDIIGEVLIPLSGIELSGKMLMNEIKRVKRSGR--G 289
   | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 328 DICEFLRYVPTAGKLVITLAKNLKMKDVGISDPYVXIAIMQNGKRLKKTYSVKCT 387
OY 350 PNAVFNELEVFDPCEGLDISVEFLVDSERGRREVIGOLVGAAGGTGGEHKKEIC 409
   | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 388 LNPYNESEFSEYFEPQOMOKICLVTVVDYDRIGTSEPIGRCLGCMGTETLRHMSDM 447
OY 410 DYPFRQIARWHLCD 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 448 ASPRREIADWHTLKD 462

```

```

RESULT 4
SY65_APLCA STANDARD: PRT; 426 AA.
ID SY65_APLCA P41823;
AC P41823;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE SYNAPTOTAGMIN (P65).
OS Aplysia californica (California sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidae;
OC Aplysiidae; Aplysia.
OX NCBI_TaxID=6500;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ganglion;
RA Hu Y.;
RL Submitted (NOV-1993) to the EMBL/Genbank/DDI databases.
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND VESICLE-LIKE
CC STRUCTURES.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U03125; AAA03567.1; -.
CC HSSP, P21707; IRSY.
CC DR InterPro: IPR000008; C2.
CC DR InterPro: IPR002149; LRI.
CC DR InterPro: IPR001565; Synaptotagmin.
CC DR Pfam: PF00168; C2; 2.
CC DR PRINTS: PR00360; C2DOMAIN.
CC DR PRINTS: PR00399; SYNAPTOTAGMIN.
CC DR SMART: SM00239; C2; 2.
CC DR PROSITE: PS00499; C2_DOMAIN_1; 2.
CC DR PROSITE: PS50004; C2_DOMAIN_2; 2.

```

KW Transmembrane; Repeat; Synapse; Glycoprotein.  
 FT DOMAIN 1 67 VESICULAR (POTENTIAL).  
 FT TRANSSEM 67 92 POTENTIAL.  
 FT DOMAIN 93 426 CYTOSOLIC (POTENTIAL).  
 FT DOMAIN 145 393 PHOSPHOLIPID BINDING (PROBABLE).  
 FT DOMAIN 167 254 C2 DOMAIN 1.  
 FT DOMAIN 298 389 C2 DOMAIN 2.  
 SO SEQUENCE 426 AA; 47459 MW; 2570853DEDEA85 CRC64;

Query Match 27.6%; Score 606.5; DB 1; Length 426;  
 Best Local Similarity 36.0%; Pred. No. 1.9e-34;  
 Matches 151; Conservative 72; Mismatches 126; Indels 71; Gaps 13;

QY 10 EFDEIP--VVGIFSAFGVFTVSLFA--CORSSSKNKTPEYKGVHVKGVDTYPE 65  
 DB 64 ETEPLTMAVLIILAGSLFLVCCVYCVCRSKRKKRKKKGGK-----LKGA---V 112  
 QY 66 NLNSKKKFGADKNEVKNRPVAPKNSLHLEKRDJLNGNPFKTNLKPGSPDLENATPKL 125  
 DB 113 DLKSVQLLG--NSYKKEP-----DDELPLVN-----MEDNDA---- 143  
 QY 126 FLEGEKESVPSLSKSTSLTSEKOEKLTFLFSLEYNEFKRAVFNINKEARGLPAMDE 185  
 DB 144 -----ESTKS-----EVKIGKLOFSLDYDFOKGLSVNYIQAADLPGM- 182  
 QY 186 QSMSTDPYIKMTLPEKKHKVTRVLRKTLDPAFDETFTFYGIPTYQIOELALHFTLSF 245  
 DB 183 MSGSDPYVYVLLPDKKKKYEKTKHRTLNPNVNESTFP-KVYAVAGSKILTFVYDF 241  
 QY 246 DRESRDIIIEVLIPLSGIESEGKMLNREIK--RNVKSSSGRELLISCYOSTN 302  
 DB 242 DRESKHQIOVOVPLNSIDL--GRVVEDMKDLOSPRESEKREKLDFCSRYVPTAG 299  
 QY 303 TLVYVYKARHLKRPDVSGLSDPYVKNVYHAKKRISKTKHYKCPNVPFELVFDI 362  
 DB 300 KLTVVILEAKNLKMDVGLSDPYVKIALLOGTKRLKKKTKTKTLNPNFNESEGEV 359  
 QY 363 PCSELEIDISVEFLVDSERGSRENYIGOLVIGAAEGSGHMKELICYDPRRIAKMYL 422  
 DB 360 PFEGIOKVTLIIIVVDYDRIIGTSEPIGRCVIGCNSSGTELRLHNSMDLANRRRIAQMHL 419

RESULT 5  
 SYTL2 MOUSE  
 AC SYTL2 MOUSE STANDARD: PRT: 422 AA.  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE STNAPOTAGMIN II (SYTLII).  
 GN SYTL2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI-TaxId=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=95050743; PubMed=7961887;  
 RA Fukuda M., Aruga J., Nishibe M., Aimoto S., Mikoshiba K.;  
 RT "Insitol-1,3,4,5-tetrakisphosphate binding to C2B domain of  
 RT IP4BP/synaplotagmin II";  
 RL J. Biol. Chem. 269:29206-29211(1994).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RA Adachi R., Telch A.H., Nigam R.;  
 RT "Genomic structure of the murine Sytl2 gene";  
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBP databases.  
 CC -!- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS  
 CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE  
 CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT  
 CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL  
 CC BACKBONE.

CC -!- SUBUNIT: HOMOTETRAMER (PROBABLE).  
 CC -!- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.  
 CC -!- DOMAIN: THE FIRST C2 DOMAIN MEDIATES CA(2+)-DEPENDENT PHOSPHOLIPID  
 CC BINDING.  
 CC -!- SIMILARITY: CONTAINS 2 C2 DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D37793; BA007041.1; -.  
 DR EMBL: AF257303; AA68987.1; -.  
 DR HSSP: P21707; IRSY.  
 DR MGD: M61:99666; Sytl2.  
 DR InterPro: IPR000008; C2.  
 DR InterPro: IPR002149; LRI.  
 DR InterPro: IPR001565; Synaplotagmin.  
 DR Pfam: PF00168; C2; 2.  
 DR PRINTS: PR00360; C2DOMAIN.  
 DR PRINTS: PR00399; SYNAPTOTAGMIN.  
 DR SMART: SM00239; C2; 2.  
 DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.  
 DR PROSITE: PS00004; C2\_DOMAIN\_2; 2.  
 KW Transmembrane; Repeat; Synapse; Glycoprotein.  
 FT TRANSSEM 67 92 VESICULAR (POTENTIAL).  
 FT DOMAIN 93 426 CYTOSOLIC (POTENTIAL).  
 FT DOMAIN 145 382 PHOSPHOLIPID BINDING (PROBABLE).  
 FT DOMAIN 156 245 C2 DOMAIN 1.  
 FT DOMAIN 287 378 C2 DOMAIN 2.  
 FT CAROHD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 422 AA; 47262 MW; B4BD13FF70E0A81B CRC64;

Query Match 27.5%; Score 604.5; DB 1; Length 422;  
 Best Local Similarity 35.6%; Pred. No. 2.6e-34;  
 Matches 151; Conservative 63; Mismatches 127; Indels 83; Gaps 12;

QY 10 EFDEIP--VVGIFSAFGVFTVSLFA--WICQSSSKNKTPEYKGVHVKGVDTI 62  
 DB 57 ETKIPLPFWALIAVAVAGLLITCCFCICCKCCCKKKKKKKG-----KGM-- 105  
 QY 63 YPENLNSKKKFGADKNEVKNRPVAPKNSLHLEKRDJLNGNPFKTNLKPGSPDLENAT 122  
 DB 106 -KNANMKDMKGGQDDDDA-----ETGLGEG----- 130  
 QY 123 PRLFGEKESVPSLSKSTSLTSEKOEKLTFLFSLEYNEFKRAVFNINKEARGLP 182  
 DB 131 -----EGGE-----EKEPENLGLKLOFSLDYFQANOLTVGLAAELPA 171  
 QY 183 MDEQMSDPIYIKMTLPEKKHKVTRVLRKTLDPAFDETFTFYGIPTYQIOELALHFTI 242  
 DB 172 LD-MGSTDPIYVYVLLPDKKKKYEKTKHRTLNPNVNESTFP-KVYQOELAGTLVMAI 229  
 QY 243 LSFDFRSDDIIEVLIPLSGIEL---SEGKMLNREIKKNVKKSSGREGELISCYQ 298  
 DB 230 YQDFRFSKHDIIIEGVKVMYNTVDGQPIEMRDLQGE--KKEPEK--LADICTSRVY 284  
 QY 299 STTNLTGVVYKARHLKRPDVSGLSDPYVKNVYHAKKRISKTKHYKCPNVPFELF 358  
 DB 285 PTAGLTVLCIEAKNLKMDVGLSDPYVNIHLMQNGKRLKKKTKTVKKTLPNFNESEF 344  
 QY 359 VPDICELEIDISVEFLVDSERGSRENYIGOLVIGAAEGSGHMKELICYDPRRIAK 418  
 DB 345 SEIIFPDIOKVYVYVLDYDKLGNKALGKIFGVSNTGTGLRLHNSMDLANRRRIAQ 404  
 QY 419 WHVL 422  
 || |



DB 405 WHSL 408

RESULT 6

SYT2\_RAT STANDARD: PRT; 422 AA.

AC P29101; 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DE 20-AUG-2001 (Rel. 40, Last annotation update)

SYNAPTOTAGMIN II (SYTII).

SYT2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI\_TaxId=10116;

OX NCB1\_TaxId=10116;

RP SEQUENCE FROM N.A.

RA MEDLINE=91310620; PubMed=1856191;

RX Geppert M., Archer B.T. III, Suedhof T.C.; "Synaptotagmin II. A novel differentially distributed form of synaptotagmin.";

RL J. Biol. Chem. 266:13548-13552(1991).

CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL BACKBONE.

CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).

CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.

CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN PHYSIOGENETICALLY OLDER BRAIN REGIONS SUCH AS THE SPINAL CORD, BRAIN STEM AND CEREBELLUM.

CC -1- DOMAIN: THE FIRST C2 DOMAIN MEDIATES CA(2+)-DEPENDENT PHOSPHOLIPID BINDING.

CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.

CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M64488; AAA63502.1; -

DR PIR: A39454; BMR2Y.

DR HSSP: P21707; IRSY.

DR InterPro: IPR000008; C2.

DR InterPro: IPR002149; LRI.

DR InterPro: IPR001565; Synaptotagmin.

DR Pfam: PF00168; C2; 2.

DR PRINTS: PR00360; C2DOMAIN.

DR SMART: SM00399; SYNAPTOTAGMIN.

DR SMART: SM00239; C2; 2.

DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.

DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.

KW Transmembrane; Repeat; Synapse; Glycoprotein.

KW DOMAIN 1

FT TRANSMEM 61 60 VESICULAR (POTENTIAL).

FT TRANSMEM 61 87 POTENTIAL.

FT DOMAIN 88 422 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 136 382 PHOSPHOLIPID BINDING (PROBABLE).

FT DOMAIN 156 245 C2 DOMAIN 1.

FT DOMAIN 287 378 C2 DOMAIN 2.

FT CAROHND 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 422 AA; 47209 MW; D852AF5387E0C7FD CRC64;

Query Match 27.4%; Score 603.5; DB 1; Length 422;

Best Local Similarity 35.6%; Pred. No. 3e-34;

Matches 151; Conservative 63; Mismatches 127; Indels 83; Gaps 12;

OY 10 EFDEIP-----TVVGIARSAGLVFTVSLFA---WICCRSSKSKNKTIPPYKVFVHLKGVDI 62

DB 57 EINKIPLPKALITAMAVAGLLLTCCFCICCKKCCCKKKKKKEG-----KGM-- 105

OY 63 YPENLSKKRKFAGADKNEVKNKPAPVKNLSLHDLERDLGNFPKTNLRGSPSDEMAT 122

DB 106 -KNAAMNKDMKGGDDDA-----ETGLFEG----- 130

OY 123 PKFLFEGEKESVPESLSKSTSLTSEKQEKLTLPFSLYENFERKAPFVNIEKAGLPA 182

DB 131 -----EGEGE-----EKEPENLKGKLFSDYDQANQLVGVYQAAELPA 171

OY 183 MDQSWTSPDYIKMTLPEKKHKVKTFRVLKTPDPADEFETFGYIPYQIQELALHFTI 242

DB 172 LD-MGSTDSPYVAVFLPDKKKYETVKRKTLPNAPNFTFT-KVPYQELGKTLVMAI 229

OY 243 LSPDRFSRDDIIGVLPISGIEL---SEKMLMNRITIKRNRKSSRGELLISLCYQ 298

DB 230 YDRDRFSKHDIIGEVAVPMVTVDLGPIEBWRDLQGE--KEPER--LGDICTSLRY 284

OY 299 STTNITLVVYVLRKRLPKSDVSGLSDPYKVNLYHAKRKSKTKVKKCTPNAVNELF 358

DB 285 PTGKLTVCLEKKNLKKMDVGLSDPYVYIHLONGKRLKKTKTYKKTLNPFNESEF 344

OY 359 VPDIPCEGLDIEVFVLVDSEKRSREVIQGLVLAAGTGEHMKETCDVPRQIAK 418

DB 345 SEETPEQIKVQVYVTVLDYDKLGKNEALGKIFVGSNATGELRHMSDLNPRPIAO 404

OY 419 WHVL 422

DB 405 WHSL 408

RESULT 7

SY63\_DISOM STANDARD: PRT; 537 AA.

ID SY63\_DISOM P24507;

AC 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE SYNAPTOTAGMIN C (SYNAPTIC VESICLE PROTEIN O-P65-C).

GN P65-C.

OS Discopoge ommata (Electric ray).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; Elasmobranchii; Squalae; Hypnosquales; Pristiogelae; Batoidae; Torpediniformes; Narcinoidae; Narcinidae; Discopoge.

OC NCB1\_TaxId=7785;

OX NCB1\_TaxId=7785;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=91273991; PubMed=2054189;

RX Wendland B., Miller K.G., Schilling J., Scheller R.H.; "Differential expression of the p65 gene family.";

RT Neuron 6:993-1007(1991).

CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL BACKBONE.

CC -1- SUBUNIT: HOMODIMER OR HOMOTRIMER (POSSIBLE).

CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES IN NEURONS.

CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.

CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M64277; AAA49229.1; -

DR PIR: JH0415; JH0415.

DR HSSP: P21707; 1RSY.  
 DR Interpro: IPR000008; C2.  
 DR Interpro: IPR002149; LRI.  
 DR Interpro: IPR001565; Synaptotagmin.  
 DR Pfam: PF00168; C2\_2.  
 DR PRINTS: PR00360; C2DOMAIN.  
 DR PRINTS: PR00399; SYNAPTOTAGMN.  
 DR SMART: SM00399; C2\_2.  
 DR PROSITE: PS00459; C2\_DOMAIN\_1; 2.  
 DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.  
 DR Transmembrane; Repeat; Synapse; Multigene family.  
 DR DOMAIN 1  
 FT TRANSMEM 53 78 VESICULAR (POTENTIAL).  
 FT DOMAIN 79 537 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 228 477 PHOSPHOLIPID BINDING (PROBABLE).  
 FT DOMAIN 252 339 C2 DOMAIN 1.  
 FT DOMAIN 382 473 C2 DOMAIN 2.  
 SQ SEQUENCE 537 AA; 61300 MW; 2792F910CFBCE682 CRC64;

Query Match 27.0%; Score 594.5; DB 1; Length 537;  
 Best Local Similarity 31.5%; Pred. No. 1,7e-33;  
 Matches 156; Conservative 88; Mismatches 140; Indels 111; Gaps 19;

OY 6 TSREPEPEITVVGIFSAF-GVFT-VSIF-AMTC-----QKSSSKNTPTPY 51  
 DB 46 TRHLEIDISLSVLYTFCGYLVGLVSWKLCWIPWMDKGLNFORSDQH---FH 102  
 OY 52 KFYH-----VLKGVDTYPENLNSKKKRGADKNEVKNKPAVPKNSLHDLERDLN 102  
 DB 103 QHLHHHSHTDLTVRVDGPE-----MPERS-YLDLE----- 135  
 OY 103 GNEPKTKMK-----PGSPSDL-----EN-----A 121  
 DB 136 -SYGESIKIKLSQTSPIPDVTSSGSKENNIPNAHSQOVSAAPPATRFNSLPPIPOQS 194  
 OY 122 TPKLFLGEKE-----SVSPESLKSSTLYSEKOEKL--GTLFFSLDENVFERKAFV 171  
 DB 195 SPERGQADKQVQVNSIGQIKPELVKQKSIDTEAKKQKQKNCGRINFMRYTTEQLV 254  
 OY 172 VNIKEAGGLPAMDEQMSDPIYKMTILPEKKHKVKTFLVKTLDPAFDETFEYGIPT 231  
 DB 255 VKIKKALDLPAKDAANGF-SDPYVKIYLLPDRKKKQFQVHKRKLINPFEETQF-NVPE 312  
 OY 232 QIOELAHFLLSEFDRSDDIIGVLLP--LSGIEISEGMMNRRITIKRNRKSSGSG 289  
 DB 313 ELQNRKLFHFVDFDRSHDLIGOVVDNLFEFSDSEDTTW-RDILEATSEKAD-LG 370  
 OY 290 ELLISLCYOSTTNTLVVYVVKARHLPKSDVSGSLDPYKYNLHYAKKRISKKTTHYKCT 349  
 DB 371 EINESLCTLEFAGRLITITIKATNKAMDLGFSDPYKSLCIDERRLKRTSTTKMT 430  
 OY 350 PNAVENLEVEDIPCEGLDIEVEFLVDSEGRSRENYIQOLVGAAGTGGEHKKEIC 409  
 DB 431 LNPVYNALVDFDIPNEMHVNIIAVMDYDCIGHNEVIGMCRVGNATDGPGRHNEML 490  
 OY 410 DYPRROIAKHWLDCD 424  
 DB 491 ANPRKPIEQWHQIE 505

RESULT 8  
 SYTL CAEEL  
 ID SYTL CAEEL STANDARD: PRT: 441 AA.  
 AC P34693;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE SYNAPTOTAGMIN 1  
 GN SMT-1 OR F31E8.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI\_TaxID=6239;  
 RP (1)  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE-93113960; PubMed-8391930;  
 RA Nonet M.L., Grundahl K., Meyer B.J., Rand J.B.;  
 RT "Synaptic function is impaired but not eliminated in C. elegans  
 RL mutants lacking synaptotagmin.";  
 RN Cell 73:1291-1305(1993).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Du Z.;  
 RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS  
 CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE  
 CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT  
 CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL  
 CC BACKBONE (BY SIMILARITY).  
 CC -1- SUPRACELLULAR LOCATION: SYNAPTIC VESICLES AND VESICLE-LIKE  
 CC STRUCTURES.  
 CC -1- TISSUE SPECIFICITY: LOCALIZED TO REGIONS KNOWN TO BE RICH IN  
 CC SYNAPSES AND APPEARS TO BE ASSOCIATED WITH SYNAPTIC VESICLES.  
 CC ALSO FOUND IN SOME NONNEURONAL SECRETORY STRUCTURES.  
 CC -1- DISEASE: MUTANTS EXHIBIT SEVERE BEHAVIORAL ABNORMALITIES THAT ARE  
 CC CHARACTERISTIC OF DEFICIENCIES IN SYNAPTIC FUNCTION, INCLUDING  
 CC SEVERE LOCOMOTION, FEEDING, AND DEFECATION DEFECTS.  
 CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: L15302; AAA28145.1; -;  
 DR EMBL: U55856; AAA98023.1; -;  
 DR PIR: A40707; A40707.  
 DR HSSP: P21707; 1RSY.  
 DR Wormpep: F31E8.2; CR02711.  
 DR Interpro: IPR000008; C2.  
 DR Interpro: IPR002149; LRI.  
 DR Interpro: IPR001565; Synaptotagmin.  
 DR Pfam: PF00168; C2\_2.  
 DR PRINTS: PR00360; C2DOMAIN.  
 DR PRINTS: PR00399; SYNAPTOTAGMN.  
 DR SMART: SM00239; C2\_2.  
 DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.  
 DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.  
 DR Transmembrane; Repeat; Synapse; Glycoprotein.  
 FT TRANSMEM 1 69 VESICULAR (POTENTIAL).  
 FT DOMAIN 70 96 POTENTIAL.  
 FT DOMAIN 97 441 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 175 262 C2 DOMAIN 1.  
 FT DOMAIN 306 397 C2 DOMAIN 2.  
 SQ SEQUENCE 441 AA; 49904 MW; F8D174337BB472DB CRC64;

Query Match 27.0%; Score 594; DB 1; Length 441;  
 Best Local Similarity 40.2%; Pred. No. 1.4e-33;  
 Matches 148; Conservative 57; Mismatches 111; Indels 52; Gaps 14;

OY 70 KKKRGADKNEVKNKPAVPKNSLHDLERDLNNGNPK-----TNLKPSPSDLENAT 122  
 DB 97 KKLFGKKRNGE-KMK-----KGLKGFGKGQGVYDGKNTQ-GMAODLEELG 141  
 OY 123 PKLFLGEKESVSPESLKSSTLSFSEKQD-KLGTLPFSLEYNFERKAFVNIKEANGLP 181  
 DB 142 DAM-RQNEKE-----AEEKEVKGIRIQYKLDYDFQGGTLVTVIQAEDLP 187



```
CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.
CC -1- DOMAIN: THE FIRST C2 DOMAIN MEDIATES CA(2+)-DEPENDENT PHOSPHOLIPID
CC BINDING.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L05922; AAA87360.1; .
CC HSSP: P21707; IRSY.
CC Interpro: IPR000008; C2.
CC Interpro: IPR002149; LRI.
CC Interpro: IPR001565; Synaptotagmin.
CC Pfam: PF00168; C2; 2.
CC PRINTS: PR00360; C2DOMAIN.
CC SMART: SM00239; C2; 2.
CC PROSITE: PS00499; C2_DOMAIN_1; 2.
CC PROSITE: PS50004; C2_DOMAIN_2; 2.
CC Transmembrane: Repeat: Synapse; Glycoprotein.
CC FT DOMAIN 1 53 VESICULAR.
CC FT TRANSMEM 54 80 POTENTIAL.
CC FT DOMAIN 136 422 CYTOPLASMIC.
CC FT DOMAIN 157 245 PHOSPHOLIPID BINDING (PROBABLE).
CC FT DOMAIN 287 378 C2 DOMAIN 1.
CC FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 422 AA; 47623 MW; C158C34DA8E56BA CRC64;

Query Match 26.4%; Score 581; DB 1; Length 422;
Best Local Similarity 39.9%; Pred. No. 1e-32;
Matches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;

OY 105 FPKTNLKPGSPDLENATPKLFLEGEKESVPESLK---SSTSLT-SEK-----QKLG 155
DB 85 FKKNNKKKKGKGGKGNAINMKDVKDGLTKMDQALKDDAETGLDGEKEEPKEEKL 144
OY 156 TLFPSLEYNEFKKAFVNIKEKRGKLPAMDQSWTSDPYIKMTILPEKKHKVTRVLKRTL 215
DB 145 KLOSLDLDYDFONNQLVIGITQAEIPLAD-MGTSIDPYKVFLLPDKKKKFEKVKHRTKL 203
OY 216 DPADFETFFYGIPTQIOELAHFTLISDFRSRDDIIGEVLLPLSGIEL--SEGM 271
DB 204 NPFVNEQDFTF-KVPSVSLGKTLVMAVVDPRFSKHDIIEFKVPMATVDFGHVEEMWD 262
OY 272 LAMREIIRKRVKSSRGCELLISICYSTNTLTIVVYLKARHLPKSDVSGLSDPYKVN 331
DB 263 LQSAE--KEQEK--LGDICFSLRKVPYTAGKLTIVVLEKRNKKMDYGLSDPYKIH 317
OY 332 YHAKKRISKKTHVKKCPNNAFVNELFVFDIPCGELDISVEFLVDSEKSGNEVIGQL 391
DB 318 MONGKRIKKKTKTIKKLNPIYVNESFSEVFPDIOKVQVVVTVLDYDKIGNDAIGKV 377
OY 392 VLGAAGETGGEHMKELIDYPRQIAKMHVL 422
DB 378 FVGYNSTGAEIARHMSDMLANPRPIAOWHTL 408

RESULT 11
SYTL_HUMAN
ID SYTL_HUMAN STANDARD: PRT; 422 AA.
AC P21579;
DT 01-MAY-1991 (rel. 18, Created)
DT 01-MAY-1991 (rel. 18, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
```

```
DE SYNAPTOTAGMIN I (SYTL) (p65).
GN SYTL OR SYTL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX SEQUENCE FROM N.A.
RX MEDLINE=91093190; PubMed=1840599;
RA perin M.S., Johnston P.A., Oezcelik T., Jahn R., Francke U.,
RA Suedhof T.C.;
RT "Structural and functional conservation of synaptotagmin (p65) in
RT Drosophila and humans."
RL J. Biol. Chem. 266:615-622(1991).
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE. A CA2+-DEPENDENT INTERACTION BETWEEN SYNAPTOTAGMIN AND
CC PUTATIVE RECEPTORS FOR ACTIVATED PROTEIN KINASE C HAS ALSO BEEN
CC REPORTED. IT CAN BIND TO AT LEAST THREE ADDITIONAL PROTEINS IN A
CC CA2+-INDEPENDENT MANNER. THESE ARE NEUREXINS, SYNTAXIN AND AP2.
CC -1- SUBUNIT: HOMOTETRAMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.
CC -1- DOMAIN: THE FIRST C2 DOMAIN MEDIATES CA(2+)-DEPENDENT PHOSPHOLIPID
CC BINDING.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M5047; AAA60609.1; .
CC PIR: A39052; BMR017.
CC HSSP: P21707; IRSY.
CC MIM: 183605; .
CC Interpro: IPR000008; C2.
CC Interpro: IPR002149; LRI.
CC Interpro: IPR001565; Synaptotagmin.
CC Pfam: PF00168; C2; 2.
CC PRINTS: PR00360; C2DOMAIN.
CC SMART: SM00239; C2; 2.
CC PROSITE: PS00499; C2_DOMAIN_1; 2.
CC PROSITE: PS50004; C2_DOMAIN_2; 2.
CC Transmembrane: Repeat: Synapse; Glycoprotein.
CC FT DOMAIN 1 53 VESICULAR.
CC FT TRANSMEM 54 80 POTENTIAL.
CC FT DOMAIN 81 422 CYTOPLASMIC.
CC FT DOMAIN 136 382 PHOSPHOLIPID BINDING (PROBABLE).
CC FT DOMAIN 157 245 C2 DOMAIN 1.
CC FT DOMAIN 287 378 C2 DOMAIN 2.
CC FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 422 AA; 47573 MW; 467F7C58E411AFA9 CRC64;

Query Match 26.4%; Score 581; DB 1; Length 422;
Best Local Similarity 39.9%; Pred. No. 1e-32;
Matches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;

OY 105 FPKTNLKPGSPDLENATPKLFLEGEKESVPESLK---SSTSLT-SEK-----QKLG 155
DB 85 FKKNNKKKKGKGGKGNAINMKDVKDGLTKMDQALKDDAETGLDGEKEEPKEEKL 144
OY 156 TLFPSLEYNEFKKAFVNIKEKRGKLPAMDQSWTSDPYIKMTILPEKKHKVTRVLKRTL 215
DB 145 KLOSLDLDYDFONNQLVIGITQAEIPLAD-MGTSIDPYKVFLLPDKKKKFEKVKHRTKL 203
```

CC	modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a> or send an email to <a href="mailto:license@isb.sib.ch">license@isb.sib.ch</a> ).
CC	EMBL; X52772; CAA36981.1; ..
DR	PIR; S09595; S09595.
DR	PDB; 1RSY; 08-MAY-95.
DR	PDB; 1BYN; 21-OCT-98.
DR	InterPro; IPR000008; C2.
DR	InterPro; IPR002149; LRI.
DR	InterPro; IPR001565; Synaptotagmin.
DR	Pfam; PF00168; C2; 2.
DR	PRINTS; PR00360; C2DOMAIN.
DR	PRINTS; PR00399; SYNAPTOTAGMN.
DR	SMART; SMO0239; C2; 2.
DR	PROSITE; PS00499; C2_DOMAIN_1; 2.
DR	PROSITE; PS00004; C2_DOMAIN_2; 2.
KW	Transmembrane; Repeat; Synapse; Glycoprotein; 3d-structure.
FT	DOMAIN 1 53
FT	TRANSMEM 53 79
FT	POTENTIAL.
FT	DOMAIN 80 421
FT	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 135 381
FT	PHOSPHOLIPID BINDING (PROBABLE).
FT	DOMAIN 156 244
FT	C2 DOMAIN 1.
FT	DOMAIN 286 382
FT	C2 DOMAIN 2.
FT	CARBOHD 24 24
FT	N-LINKED (GLCNAC. . .).
SO	SEQUENCE 421 AA; 47453 MW; FB78A1C392EAD7E2 CRC64;
Query Match	26.4%; Score 580; DB 1; Length 421;
Best Local Similarity	40.2%; Pred. No. 1.2e-32;
Matches 133; Conservative	63; Mismatches 115; Indels 20; Gaps 8;
OY	105 FPKTKLPKPGSPEDLNAPKLFLEGEKESVSPESLK---SPTSLT-SEEK---QEKLG 155
DB	84 FKKKKKKKKKKKKKKGNALNMKDKVKGDKLGTMDQALKDDAEGLDGEKEBKPEKKEKLG 143
OY	156 TLFSLEYNEFKKAFVNIKEARGLPRAMDEGSMISDPYIKMTLPKKNKKVTRVLRKTL 215
DB	144 KLYQSLDVPQFNQNLVGLIQAEALPALD-MGQTSDPYVKVLLPKEKKKKFETKVNRKTL 202
OY	216 DPAPDETTFPGIPRYQQLALHTLTLSDFRFSDDLTIGVLRPLSLIEL---SECKM 271
DB	203 NPFVNEQGLTF-KVPSLELGKTLVNAVYVDFRFSKNDLTIGEEKVPMNVVDGHTVEEMRD 261
OY	272 LNMREITKRNKVSRSRGELLISLCYOSTTNTLVVAVKARLHPKSDVSGSLSDPYKVNLT 331
DB	262 LOSAE--KEQEK---LDGICFSLKYVTPAGKLTIVIIIEAKKLKMDVGGSLDPYVKIHL 316
OY	332 YHAKKRISKKTTHVKCTPRNAVFENLFVEDIPCEGLIEDISVEFLVLDSEGRSNEVIGQL 391
DB	317 MONGRKLKKKKTKTKKTLNPNYNSFSFEVDFEQIQKVQVYVVLVDYDKIGKNDALDKV 376
OY	392 VIGAAAEGTGGEHMKELCDYPRKQIAKKNHV 422
DB	377 FVGYNSTGAELRHMSDIILANPRRPIAQMHTL 407
RESULT 13	
ST61_DISOM	STANDARD; PRT; 427 AA.
AC	P24505;
DT	01-MAR-1992 (Rel. 21, Created)
DT	01-MAR-1992 (Rel. 21, Last sequence update)
DT	01-NOV-1995 (Rel. 32, Last annotation update)
DE	SYNAPTOTAGMIN A (SYNAPTIC VESICLE PROTEIN O-P65-A).
GN	P65-A.
OS	Discopoge ommata (Electric Ray).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC	Elasmobranchii; Squalea; Hypnosqualea; Pristioralea; Batoidae;
OC	Torpediniformes; Narchinoidei; Narchinae; Discopoge.
OX	NCBI_TaxID=7785;
RP	SEQUENCE FROM N.A.

RA MEDLINE-91273991; PubMed-2054189;  
RA "Differential expression of the p53 gene family.";  
RL Neuron 6:993-1007(1991).  
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS  
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE  
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT  
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL  
CC BACKBONE.  
CC -1- SUBUNIT: HOMODIMER OR HOMOTRIMER (POSSIBLE).  
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES IN NEURONS.  
CC -1- TISSUE SPECIFICITY: FOREBRAIN, CEREBELLUM, AND NEUROENDOCRINE  
CC CELLS.  
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M64275; AAA49227.1; -  
CC PIR: JH0413; JH0413.  
CC HSSP: P21707; IRSY.  
CC InterPro: IPR000008; C2.  
CC InterPro: IPR002149; LRI.  
CC InterPro: IPR001565; Synaptotagmin.  
CC DR Pfam: PF00168; C2; 2.  
CC DR PRINTS: PRO0399; SYNAPTOTAGMN.  
CC DR SMART: SM00239; C2; 2.  
CC DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.  
CC DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.  
CC KW Transmembrane; Repeat; Synapse; Multigene family; Glycoprotein.  
CC FT DOMAIN 1 57 VESICULAR (POTENTIAL).  
CC FT TRANSMEM 58 84 POTENTIAL.  
CC FT DOMAIN 85 427 CYTOPLASMIC (POTENTIAL).  
CC FT DOMAIN 141 387 PHOSPHOLIPID BINDING (PROBABLE).  
CC FT DOMAIN 162 250 C2 DOMAIN 1.  
CC FT DOMAIN 292 383 C2 DOMAIN 2.  
CC FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC SEQUENCE 427 AA; 47768 MW; D23F9488A8A600A CRC64;  
SO  
Query Match 26.3%; Score 579.5; DB 1; Length 427;  
Best Local Similarity 38.8%; Pred. No. 1.3e-32;  
Matches 128; Conservative 66; Mismatches 115; Indels 21; Gaps 7;  
QY 107 KYNKPGSPDLENA-----TKLLEBKESVSPESIKSSTSLTS-----EKKOKLGT 156  
DB 91 KNNKKKKEKGGKNAATMDVKEMKSKSEQALKDEDEATGTGDKKEEEDKRLK 150  
QY 157 LFTSLEYNEERKAVVNIKEARGLPAMDQSTSDPYIKMTLPEKKHKVTVLKRKTL 216  
DB 151 LQESLDLDFQNNOLVIGIIQAAELPALDVGG-TSDPYVYKVVLPDKKKKYEKTVHRRKTLN 209  
QY 217 PARDFTFTFYGYPIYTOIQELALHTLILSFDRFSRDDIGEVLLPLSGIEL---SEGKML 272  
DB 210 PVNESRIF-KIYSELGKTLVMAVYDFDRFSKHDVGEAKVPMNTVDEGHVTEMRDL 268  
QY 273 MNEEIIIRNRKSSGSGELLISLCYOSTNTNLJVVLKARLPLPSDVGSLDPYVNVNY 332  
DB 269 QGAEK-KEDEK---LGDICFLRYVPTAGKLVYLLEAKNLKMDVGGSLDPYVNIHM 323  
QY 333 HAKKRISKTKTHYKCTPAVNEFLVFDICDEGEDISEVFLVDSERSRSREVIGOLY 392  
DB 324 QNKRRLKKKTKTITKNTLMPYNNESFSFEVFEQIQVQVYVLYLDKIGKNDALGRKF 383  
QY 393 LGAAAGTGTGHHKELCDYPRQIAKHVYL 422

DB 384 VGYNSTAEELRHWSMDMLANPRRPIAOWHTL 413  
RESULT 14  
SYTL\_CHICK STANDARD; PRT; 424 AA.  
ID P47191;  
AC 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 01-NOV-1995 (Rel. 32, Last annotation update)  
OS SYNAPTOTAGMIN I (P65).  
OC Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-93374184; PubMed-8365570;  
RA Lou X., Bixby J.L.;  
RT "Coordinate and noncoordinate regulation of synaptic vesicle protein  
RT genes during embryonic development.";  
RL Dev. Biol. 159:327-337(1993).  
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS  
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE  
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT  
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL  
CC BACKBONE.  
CC -1- SUBUNIT: HOMOTRIMER (PROBABLE).  
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.  
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: S64957; AA28081.1; -  
CC HSSP: P21707; IRSY.  
CC InterPro: IPR000008; C2.  
CC InterPro: IPR002149; LRI.  
CC InterPro: IPR001565; Synaptotagmin.  
CC DR Pfam: PF00168; C2; 2.  
CC DR PRINTS: PRO0399; SYNAPTOTAGMN.  
CC DR SMART: SM00239; C2; 2.  
CC DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.  
CC DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.  
CC KW Transmembrane; Repeat; Synapse; Glycoprotein.  
CC FT DOMAIN 1 55 VESICULAR.  
CC FT TRANSMEM 56 82 POTENTIAL.  
CC FT DOMAIN 83 424 CYTOPLASMIC.  
CC FT DOMAIN 138 384 PHOSPHOLIPID BINDING (PROBABLE).  
CC FT DOMAIN 159 247 C2 DOMAIN 1.  
CC FT DOMAIN 289 380 C2 DOMAIN 2.  
CC FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC SEQUENCE 424 AA; 47505 MW; C602676F8F679718 CRC64;  
SO  
Query Match 26.1%; Score 575; DB 1; Length 424;  
Best Local Similarity 37.0%; Pred. No. 2.7e-32;  
Matches 141; Conservative 71; Mismatches 129; Indels 40; Gaps 11;  
QY 68 NSKKRGAGDDKNEKKNKPAVKNSLHD-----LEKRDNGNPPKNTKPGS 114  
DB 44 NKKKF-----NKLKTKIPLPMAIATAIVAVLLITCCFLCKKCL---FKKNNKKK 96  
QY 115 PSDLENAATPKLLEGEKESVSPESLK---SSTSLT-SEKKQ-----EKLGLTFTSLEYNF 165

```

DR 97 ERGKNAIMKDVKDKGTMKDDAKDDAETGLADGEKEEPEVEKLGKIQYSLDYDE 156
OY 166 ERKAEVNIKEARGLPAMDQESMTSDPYIKMTILPEKKHKVTRVLKRTLDPAFDETFE 225
DB 157 QNNOLLVGIQAAELPALD-MGSTSDPYKVFLLPPKKKKYETKVRKRLNPFNFQPTF 215
OY 226 YQIPYQIQELALHFTILSFDRSRDDIIGEVLLPLSGIEL---SEGKMLNREILIKRN 281
DB 216 -KVPYSELGKTLVMAVYDFDRSKHDITIGEVKVAANTVDGFHVTEWMDLQSAE--KEE 272
OY 282 VKRSSGRGLLSLCYSTNTLTLYVVKARHLPKSDVSGLSDPYKVLNLYHAKKRISK 341
DB 273 QEK---LGDICSLRVYPAGKILVILEAKNKKMDVGLSDPYKILHLMQNGKRLKK 329
OY 342 KTHVKKCPNAVFNEFVDFDICEGLDISVEFLVDSEGRSNEVIGQLVGAAGTG 401
DB 330 KTTIKNTLNPNYNESEFVEPEQIQKVQIVYVLDYDKIGNDAIGVFGVGYNTGAE 389
OY 402 GEHKEICDYPROIAMHVL 422
DB 390 LRHMSDMLANPRRPIAQWHTL 410

RESULT 15
SY62_DISOM STANDARD: PRT: 439 AA.
AC P24506:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SYNAPTOTAGMIN B (SYNAPTIC VESICLE PROTEIN O-P65-B).
GN P65-B.
OS Discopge ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalae; Hypnosqualae; Pristiogryes; Batoidae;
OC Torpediniformes; Narcinoidae; Narcinidae; Discopge.
OX NCBI_TaxId=7785;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-91273991; PubMed-2054189;
RA Wendland B., Miller K.G., Schilling J., Scheller R.H.;
RT "Differential expression of the p65 gene family.";
RL Neuron 6:993-1007(1991).
CC -1- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS
CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE
CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT
CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL
CC BACKBONE.
CC -1- SUBUNIT: HOMODIMER OR HOMOTRIMER (POSSIBLE).
CC -1- SUBCELLULAR LOCATION: SYNAPTIC VESICLES IN NEURONS.
CC -1- TISSUE SPECIFICITY: SPINAL CORD, BRAINSTEM, MIDBRAIN AND ELECTRIC
CC ORGAN.
CC -1- SIMILARITY: CONTAINS 2 C2 DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SYNAPTOTAGMIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M64276; AAA49228.1; -.
CC PIR: JH0414; JH0414.
CC HSSP: P21707; IRSY.
CC InterPro: IPR000008; C2.
CC InterPro: IPR002149; LRI.
CC InterPro: IPR001565; Synaptotagmin.
CC Pfam: PF00168; C2_2.
CC PRINTS: PR00360; C2DOMAIN.
CC PRINTS: PR00399; SYNAPTOTAGMIN.
CC SMART; SM00239; C2; 2.

```

```

DR PROSITE: P50049; C2_DOMAIN_1; 2.
DR PROSITE: P50004; C2_DOMAIN_2; 2.
KM Transmembrane; Repeat; Synapse; Multigene family; Glycoprotein.
FT DOMAIN 1 74 VESICULAR (POTENTIAL).
FT TRANSMEM 75 101 POTENTIAL.
FT DOMAIN 102 439 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 153 399 PHOSPHOLIPID BINDING (PROBABLE).
FT DOMAIN 262 262 C2 DOMAIN 1.
FT DOMAIN 304 395 C2 DOMAIN 2.
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 439 AA; 49278 MW; 2033F05F8C69F39 CRC64;

Query Match 26.1%; Score 574.5; DB 1; Length 439;
Best Local Similarity 33.4%; Pred. No. 3,1e-32;
Matches 140; Conservative 77; Mismatches 149; Indels 53; Gaps 9;

OY 44 KSNKTPPYKVFVHLKGVDIYPENLNSKKKFGADKKNEVKNKPAVPKNSLHLDLEKRDNG 103
DB 20 KTRETHPQAFVAPMATATAMP-----IDTGDNSTEAGVPBGKNDVFEKIKKEKFMN- 70
OY 104 NEPKTNLKPSPSDLENAPKLF-----EGEKESVSPSELS 141
DB 71 ELQKIPLPWALAIIVISGLLTCCLCKCKCKKKKKKKKKKKKKNDIMKDVKG 130
OY 142 S-----TSLSSE-----KQKLGTLFESLETFEKKAVVNIKEARGLPAMD 184
DB 131 SGNODDDAETGLTGEKEEKEEKEKIGKIOFSLDVFQANQLTVGIQAAELPALD 190
OY 185 EQGSDPYIKMTILPEKKHKVTRVLKRTLDPAFETFTFYQIPYQIQELALHFTILS 244
DB 191 -MGSTSDPYKVFLLDPKKKRYETKVRKRLNPFNFSEFV-KVPYQELGKTLMAVVD 248
OY 245 FDRSRDDIIGEVLLPLSGIELSEGMN-REIKRNRKSSGREGELISLCYSTNT 303
DB 249 FDRSRKHDICIGVYVLTMTKVDL--GOOLEEWRDLSEAEKEPEKLDICTSLRYVPTAGK 306
OY 304 LTVVVLKARHLPKSDVSGLSDPYKVLNLYHAKKRISKKKTHVKKCPNAVFNEFDFIP 363
DB 307 LTVCLLEAKNKKMDVGVSGSDPYKILHLLONGKRLKKKTYKRTLNPNYNESEFSEIP 366
OY 364 CEGLEDISVEFLVDSEGRSNEVIGQLVGAAGTGEGHKEICDYPROIAMHVL 422
DB 367 FEQIQVQVCYVLDYDKIGNDAIGKIPVGSNASGELRHMSDMLANPRRPIAQWHTL 425

```

Search completed: December 12, 2001, 10:57:59  
Job time: 190 sec

Fri Dec 14 10:32:19 2001

us-09-680-121-2.rsp

Page 12





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 12, 2001, 10:54:24 ; Search time 25.64 Seconds  
(without alignments)  
2424.561 Million cell updates/sec

Title: US-09-680-121-2  
Perfect score: 2201  
Sequence: 1 MAPITTSREPFDEIPVVG.....KEICDYPRQIAKWHVLCDS 425

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_17:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organella:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2201	100.0	425	4 Q9H2B2	Q9H2B2 homo sapien
2	2201	100.0	426	4 Q9P2K4	Q9P2K4 homo sapien
3	1167.5	53.0	430	11 Q9RON3	Q9RON3 mus musculu
4	1162	52.8	431	4 Q9BT88	Q9BT88 homo sapien
5	1156.5	52.5	430	11 Q08835	Q08835 rattus norv
6	757	34.4	474	5 Q9UB67	Q9UB67 drosophila
7	645.5	29.3	357	5 Q9GRB0	Q9GRB0 halocynthia
8	627.5	28.5	403	11 Q9RON7	Q9RON7 mus musculu
9	627.5	28.5	403	11 Q62747	Q62747 rattus norv
10	627	28.5	424	5 Q25393	Q25393 loligo peal
11	622.5	28.3	474	5 Q9VCG7	Q9VCG7 drosophila
12	621	28.2	403	5 Q07529	Q07529 loligo peal
13	615.5	28.0	418	4 Q43581	Q43581 homo sapien
14	615.5	28.0	520	11 Q99P37	Q99P37 rattus norv
15	615.5	28.0	523	11 Q99P36	Q99P36 rattus norv
16	615.5	28.0	567	11 Q99P35	Q99P35 rattus norv
17	615.5	28.0	611	11 Q99P34	Q99P34 rattus norv
18	615.5	28.0	643	11 Q99P33	Q99P33 rattus norv
19	615.5	28.0	687	11 Q99P38	Q99P38 rattus norv

20	613	27.9	429	5 Q9BMF3	Q9BMF3 manduca sex
21	603	27.4	511	11 Q62746	Q62746 rattus norv
22	598	27.2	426	11 Q9QUK7	Q9QUK7 mus musculu
23	598	27.2	511	11 Q9RON8	Q9RON8 mus musculu
24	592	26.9	498	11 Q08625	Q08625 rattus norv
25	590	26.8	523	11 Q9RON4	Q9RON4 mus musculu
26	588	26.7	395	5 Q9URR6	Q9URR6 dugesia jap
27	586.5	26.6	280	5 Q9V4C4	Q9V4C4 drosophila
28	580.5	26.4	386	11 Q9RON5	Q9RON5 mus musculu
29	563.5	25.6	491	11 Q9RON9	Q9RON9 mus musculu
30	547	24.9	279	11 Q62748	Q62748 rattus norv
31	544.5	24.7	390	5 P91493	P91493 caenorhabdi
32	518.5	23.6	590	4 Q9BOG1	Q9BOG1 homo sapien
33	517	23.5	587	11 P97791	P97791 mus musculu
34	474	21.5	474	11 Q62807	Q62807 rattus norv
35	464.5	21.1	474	4 Q9BSW7	Q9BSW7 homo sapien
36	463	21.0	474	4 Q9NZ18	Q9NZ18 homo sapien
37	451.5	20.5	361	4 Q43330	Q43330 homo sapien
38	430.5	19.6	412	11 P70169	P70169 mus musculu
39	428	19.4	412	4 Q14184	Q14184 homo sapien
40	427.5	19.4	412	11 P70610	P70610 rattus norv
41	415	18.9	257	5 Q20354	Q20354 caenorhabdi
42	394.5	17.9	403	11 P70611	P70611 rattus norv
43	393.5	17.9	405	11 Q35527	Q35527 mus musculu
44	390.5	17.7	315	5 Q16700	Q16700 caenorhabdi
45	384.5	17.5	400	4 Q14183	Q14183 homo sapien

ALIGNMENTS

RESULT 1	
Q9H2B2	PRELIMINARY; PRT; 425 AA.
ID Q9H2B2	
DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE SYNAPNOTAGMIN IV.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RM [1]	
RP SEQUENCE FROM N. A.	
RX MEDLINE=20538435; PubMed=10938284;	
RA Ferguson G.D., Chen X.-N., Korenberg J.R., Herschman H.R.;	
RT "The Human Synapnotagmin IV Gene Defines an Evolutionary Break Point	
RT between Syntenic Mouse and Human Chromosome Regions but Retains Ligand	
RT Inducibility and Tissue Specificity.";	
RL J. Biol. Chem. 275:36920-36926(2000).	
DR EMBL: AF299075; AAC37229.1; -.	
DR InterPro: IPR000008; C2.	
DR InterPro: IPR002149; LRI.	
DR InterPro: IPR001565; Synapnotagmin.	
DR Pfam: PF00168; C2; 2.	
DR PRINTS: PR00360; C2DOMAIN.	
DR PRINTS: PR00399; SYNAPNOTAGMIN.	
DR SMART: SM00239; C2; 2.	
DR PROSITE: PS00499; C2_DOMAIN_1; UNKNOWN_1.	
DR PROSITE: PS50004; C2_DOMAIN_2; 2.	
SQ SEQUENCE 425 AA; 47958 MW; DA3DCH175CB528D CRC64;	

Query Match 100.0%; Score 2201; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 2e-153;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAPITTSREPFDEIPVVGIFSAFGLVFTVSLFAMTCCORRKSRSKTPPYFVHLKGV 60
DB	1	MAPITTSREPFDEIPVVGIFSAFGLVFTVSLFAMTCCORRKSRSKTPPYFVHLKGV 60
QY	61	DIYPENLNSKKRFGADDKNEVKNRPAVPKNSLHLDEKRDNGNFPKTMKPGSPSDLEN 120

```

Db 61 DIVPENLNKKKKGADDDKNEVKNKPAVFNKSLHLDLEKRDNGNFPKTNLKGSPSDLEN 120
QY 121 ATPKLFLEGEKESVPESLKSSTSLTSEKOEKLGTLFESLEYNEFKKAFVNIKEARGI 180
Db 121 ATPKLFLEGEKESVPESLKSSTSLTSEKOEKLGTLFESLEYNEFKKAFVNIKEARGI 180
QY 181 PAMDEQSMSTDPYIKMTLLPEKKHKVTRVLRKTLDPAFDEFTTFYGIPTOIOELALHF 240
Db 181 PAMDEQSMSTDPYIKMTLLPEKKHKVTRVLRKTLDPAFDEFTTFYGIPTOIOELALHF 240
QY 241 TILSFDRFSRDDIIGEVLLPLSGIELSEGMKLMNREIKRNVRKSSGRGELLISLCYST 300
Db 241 TILSFDRFSRDDIIGEVLLPLSGIELSEGMKLMNREIKRNVRKSSGRGELLISLCYST 300
QY 301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNVAVNELFVF 360
Db 301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNVAVNELFVF 360
QY 361 DIPCEGLDISEDIVLDSERGSRNVEYIGOLVYGAAGTGEHMKETICYPKROIANKH 420
Db 361 DIPCEGLDISEDIVLDSERGSRNVEYIGOLVYGAAGTGEHMKETICYPKROIANKH 420
QY 421 VLCDG 425
Db 421 VLCDG 425

```

```

RESULT 2
Q9P2K4 PRELIMINARY: PRT: 426 AA.
AC O9P2K4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE KIAA1342 PROTEIN (FRAGMENT).
GN KIAA1342.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN.
RX MEDLINE-20181126; PubMed-10718198.
RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL EMBL: AB037763; BAA92580.1; -.
DR EMBL; AB037763; BAA92580.1; -.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002149; LRI.
DR InterPro: IPR001565; Synaptotagmin.
DR Pfam: PF00168; C2; 2.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGMN.
DR SMART: SM00239; C2; 2.
DR PROSITE: PS00499; C2_DOMAIN_1; 2.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
FT NON_TER
SQ SEQUENCE 426 AA; 48086 MW; CA46AF76P27A7D09 CRC64;

```

```

Query Match 100.0%; Score 2201; DB 4; Length 426;
Best Local Similarity 100.0%; Pred. No. 2e-153;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAPITTSREEDDEIPYVIGISAFGLVTVS--LEFAMICQKRSKSNKTPPYKVFVHLKGV 60
Db 2 MAPITTSREEDDEIPYVIGISAFGLVTVS--LEFAMICQKRSKSNKTPPYKVFVHLKGV 61
QY 61 DIVPENLNKKKKGADDDKNEVKNKPAVFNKSLHLDLEKRDNGNFPKTNLKGSPSDLEN 120

```

```

Db 62 DIVPENLNKKKKGADDDKNEVKNKPAVFNKSLHLDLEKRDNGNFPKTNLKGSPSDLEN 121
QY 121 ATPKLFLEGEKESVPESLKSSTSLTSEKOEKLGTLFESLEYNEFKKAFVNIKEARGI 180
Db 122 ATPKLFLEGEKESVPESLKSSTSLTSEKOEKLGTLFESLEYNEFKKAFVNIKEARGI 181
QY 181 PAMDEQSMSTDPYIKMTLLPEKKHKVTRVLRKTLDPAFDEFTTFYGIPTOIOELALHF 240
Db 182 PAMDEQSMSTDPYIKMTLLPEKKHKVTRVLRKTLDPAFDEFTTFYGIPTOIOELALHF 241
QY 241 TILSFDRFSRDDIIGEVLLPLSGIELSEGMKLMNREIKRNVRKSSGRGELLISLCYST 300
Db 242 TILSFDRFSRDDIIGEVLLPLSGIELSEGMKLMNREIKRNVRKSSGRGELLISLCYST 301
QY 301 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNVAVNELFVF 360
Db 302 TMTLVVVLKARHLPKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKCTPNVAVNELFVF 361
QY 361 DIPCEGLDISEDIVLDSERGSRNVEYIGOLVYGAAGTGEHMKETICYPKROIANKH 420
Db 362 DIPCEGLDISEDIVLDSERGSRNVEYIGOLVYGAAGTGEHMKETICYPKROIANKH 421
QY 421 VLCDG 425
Db 422 VLCDG 426

```

```

RESULT 3
Q9RON3 PRELIMINARY: PRT: 430 AA.
AC Q9RON3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SYNAPTOTAGMIN XI.
GN SY11.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=CEREBELLUM;
RX MEDLINE-20002669; PubMed-10531343;
RA Fukuda M., Kanno E., Mikoshiba K.;
RT "Conserved N-terminal cysteine motif is essential for homo- and
RT heterodimer formation of synaptotagmins III, V, VI, and X.";
RL J. Biol. Chem. 274:31421-31427(1999).
DR EMBL: AB026808; BAA85780.1; -.
DR HSSP: P21707; IRSY.
DR MGD: MGI:1859447; SY11.
DR InterPro: IPR000008; C2.
DR InterPro: IPR002149; LRI.
DR InterPro: IPR001565; Synaptotagmin.
DR Pfam: PF00168; C2; 2.
DR PRINTS: PR00360; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGMN.
DR SMART: SM00239; C2; 2.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
SQ SEQUENCE 430 AA; 48359 MW; 25E7CD5C4B4BE036 CRC64;

```

```

Query Match 53.0%; Score 1167.5; DB 11; Length 430;
Best Local Similarity 51.0%; Pred. No. 1.1e-77;
Matches 233; Conservative 70; Mismatches 93; Indels 61; Gaps 8;
QY 1 MAPITTSREEDDEIPYVIGISAFGLVTVS--LEFAMICQKRSKSNKTPPYKVFVHLKGV 58
Db 1 MAETINIRSPEDVSVVAGLIGASVLVVCVSVTFVWTCCHQDAKRRKHTPPYKFIHMLK 60
QY 59 GVDIVPENLNKKK-----FGADDKNEVKNKPAVFNKSLHLDLEKRDNGNFPKTNLKGSPSDLEN 91

```

```
DB 61 GISIYPETLSNKKKIIKVRKDCGPRRESGRGNLLINAESGLISHDKDPRGSPASCMDQ 120
QY 92 LHLDEKRDNLNMF--PKTNLKGSPSDLENATPKLFLGEEKESVSPESLAKSTSLTSE 149
DB 121 LPT---KRYGELSRSPMTSLTP-----GESKPTSP-----SSPE 152
QY 150 KOEKLGTFFLESYENFERKAFVNIKEARGLPAMDQOSTDPYIKMTLLPEKKHKVTR 209
DB 153 EHYMGLSLTFSDYDNFPPKALVVTIOEAHGLFVMDQOTGSDPYIKMTLLPEKKHKVTR 212
QY 210 VLKRTLPADDETFEFTFYGIPTQIOELAHFTLSFDRFSRDDIIGEVLIPLSGIELSEG 269
DB 213 VLKRTLPADDETFEFTFYGIPTQIOELAHFTLSFDRFSRDDIIGEVVPLAGVDPSTG 272
QY 270 KMLMRREIKRNVKSSGSGELLISLCYOSTNTNTLVVVKARHLPKSDVSGLS--DPYVK 328
DB 273 KVLQRLDIIRKRIQKICISGELQVSLSTQVAKRMVTVLKRHLPKMDITGLSGNPYK 332
QY 339 VNLVHAKKRISKKRTHVKKCPNPAVFNELFVDPICGEGLEDISVEFLVDSRSGRNEYI 388
DB 333 VVVYGRKRIAKKTHVKKCTLNPFVNESFIYDIPDLDPDISIEFLVIDFDRTKNEVY 392
QY 389 GOLVGAAGAGT--GGEHKEICDYPRROIAKMHVLC 424
DB 393 GRLLGAHSVTTSGAEHMEVCESPRKPIAKWHSLE 429

RESULT 4
QY9BT88 PRELIMINARY: PRT: 431 AA.
AC Q9BT88;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE SIMILAR TO SYNAPTOTAGMIN 11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC004291; AA04291.1; -.
SQ SEQUENCE 431 AA; 48306 MW; A7F76AD745C0FF7D CRC64;

Query Match 52.8%; Score 1162; DB 4; Length 431;
Best Local Similarity 52.4%; Pred. No. 2.8e-77;
Matches 229; Conservative 77; Mismatches 111; Indels 20; Gaps 7;
```

```
DB 294 ELQVSLSYQPVNAQRMVTVVVKARHLPKMDITGLSGNPYKVVVYGRKRIAKKTHVKK 353
QY 349 TPNAVFNELFVDPICGEGLEDISVEFLVDSRSGRNEYIGOLVIGA-AAESTGGEHKE 407
DB 354 TLNPFNESFIYDIPDLDPDISIEFLVIDFDRTKNEVGRLLIGAHSVTSGAEHME 413
QY 408 ICDYPRROIAKMHVLC 424
DB 414 VCESPRKPIAKWHSLE 430

RESULT 5
QY008835 PRELIMINARY: PRT: 430 AA.
AC 008835;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE SYNAPTOTAGMIN XI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRNIN;
RA von Poser C., Ichitchenko K., Shao X., Rizzo J., Sudhof T.C.;
RL J. Biol. Chem. 0:0-0(0).
DR EMBL: AF000423; AAB58344.1; -.
DR HSSP: P21707; IRSY.
DR Interpro: IPR000008; C2.
DR Interpro: IPR002149; IRI.
DR Interpro: IPR001565; Synapcotagmin.
DR Pfam: PF00168; C2; 2.
DR PRINTS: PR00399; SYNAPTOTAGMIN.
DR SMART: SM00239; C2; 2.
DR PROSITE: PS50004; C2_DOMAIN_2; 2.
SQ SEQUENCE 430 AA; 48268 MW; C1FA13CB9177C825 CRC64;

Query Match 52.5%; Score 1156.5; DB 11; Length 430;
Best Local Similarity 51.5%; Pred. No. 7e-77;
Matches 228; Conservative 78; Mismatches 104; Indels 33; Gaps 7;
```

RESULT	6	PRELIMINARY;	PRT;	474 AA.
DR	Q906P7			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	SYNAPTOTAGMIN IV (SYTIV PROTEIN).			
GN	SYTIV OR CG10047.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7227;			
RN	SEQUENCE FROM N.A.			
RP	MDL=99394599; PubMed=10466723;			
RC	STRAIN=BERKELER;			
RC	MDL=20196006; PubMed=10731132;			
RA	Adams N.D., Ceinalik S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Mananalis P., Davis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sitting G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,			
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Bailly R.M., Basu A., Baxendale A., Bhatnagar D., Bolshakov S.,			
RA	Besson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Boulter J., Brockstein P., Brothier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dunn-Crooks S., Dunkov B.C., Dunn P.			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glock A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,			
RA	Jatani M., Katush F., Karen G.H., Ke Z., Kennison J.D., Lai Z.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lin X.,			
RA	Liu X., Lei Y., Levitsky A.C., McLeod M.P., McPherson D.,			
RA	Lasko P., Matel B., McIntosh T.A., C., McLeod M.P., McPherson D.,			
RA	Meinkoth G., Milshina N.V., Mobarry C., Morris J., Moshire A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.S., Nixon K., Nusskern D.R., Pauley J.M.,			
RA	Palazzo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Stamps M., Skupski M.P., Smith T.,			
RA	Spieler E., Spreading A.C., Stapleton M., Strong R., Sun E.,			
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wassarman D.A., Weinstein K.C., Wu D., Yang S., Yao Q.A.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,			
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;			
RT	"The genome sequence of Drosophila melanogaster."			
RL	Science 287:2185-2195(2000).			
DR	EMBL; AF181098; AAD53186.1;			
DR	EMBL; AEO03672; AAF54037.2;			
DR	HSSP; P21707; IRYN			
DR	Protein; Fbgn0028400; SytIV.			
DR	Interpro; IPR000008; C2.			

Query Match	34.4%	Score 757	DB 5	Length 474
Best Local Similarity	38.2%	Pred. No. 1.5e-47		
Matches 179	Conservative	82	Mismatches 137	Indels 70
				Gaps 14
QY	14	ITPTVGLFSAGLVTFVYLSLFANICQ-----RKSXSKTPPYKRV-----	54	
Db	16	VALLIGLTA-----AVLSSVCLICARQRLNKKQSHDASFPQTRPRPAVNSPSSGP	71	
QY	55	-IVKAGVDTIYENLNSKKFGAGDKNEKKNPAPVK-----SLIDLEKRLNG	103	
Db	72	PHYLK-----SPTPGKQKMLSPQDOSTSPINQPVVXSEEGDPAQAHAQQ-----NG	124	
QY	104	N-----FPTNKKPSPDLFENAPKPLFLEGEK--SVSEPLKSTSLTSEKOEKLTFL	158	
Db	125	NQLYTVAGDNGKLHLSHLNSHSPVETIANGSVTTTLDHSLTNGKELVTYDQYGLGTY	184	
QY	159	FSELEFNPERRKAFVNNIKEARGLPAMDDEQSM-----SDPIKMTILPERKH	204	
Db	185	FLRLRLAERLALMVSLIRCGRLPCKGSGSGTDPITGMNGRTQAADPVPVKLOLILPRKH	244	
QY	205	KKTATVLRKLTDPDPEDTFTEFYGIPTQLOELALHFLTSLFDFSSRDIDIGEVLPISGI	264	
Db	245	KKTATVVRNTNPNPYDDEDPFYGLNNDLNNNSLHFVILSFDRYSRDVIGEVVCPRLSI	304	
QY	265	E---LSEKMLMNEIETKRNVR--KSSRGCELLSLCYOSTNTLTIVVYLAKANLEKSPDS	320	
Db	305	ELGDLSKRLSLISKEIOPRLKTRAGRGRELLSLICWQPAAGRLTYVLLKANLRPMVYT	364	
QY	321	GLSDPYVNVNIIHAKKRISKKTTHVKKCTPNAVFNELEFVDIP-CEG-----LEDISVEFL	375	
Db	365	GLADPYVVIYLLNGORIGAKKTHVKKRRLSPVFNEFPADIPDIPAGAGASLEGVSELM	424	
QY	376	VLDSEKSGRNEVIGQVLVGA--AEGTGEHMKREIDYPRROIAKMHVL	422	
Db	425	LIDMDVRYTKNEVIGLELGGPNSSYALNMHNEVCNSPQROIAEMHKL	472	
RESULT	7			
Q9GRB0	PRELIMINARY	PRT	357 AA	
AC	Q9GRB0			
DT	01-MAR-2001 (TrEMBLrel. 16, Created)			
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
OS	SYNAPTOTAGMIN			
DE	Halocynthia roretzi (Sea squirt).			
NC	Halocynthia			
OC	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;			
CC	Scolidobranchia; Pyuridae; Halocynthia.			
OX	NCBI_TaxID:7729;			
KN	SEQUENCE FROM N.A.			
RA	Katsuyama Y., Matsumoto J., Okamura Y.:			
RT	Regulation of the synaptotagmin gene in ascidian embryo."			
RL	Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.			
RL	EMBL:AB041147, BAB1864.1, --			
DR	EMBL:AB041147, BAB1864.1, --			
DR	InterPro: IPR000008; C2			
DR	InterPro: IPR002149; C1			
DR	InterPro: IPR001565; Synaptotagmin.			
DR	Frame: P00168, C2_2			
DR	Frame: P00168, C2_2			
DR	PRINTS: PR00360; C2DOMAIN			
DR	PRINTS: PR00360; C2DOMAIN			
DR	PROSITE: PS50049; C2_DOMAIN_2; 2			
DR	PROSITE: PS50049; C2_DOMAIN_2; 2			
DR	SEQUENCE 474 AA: 52229 MW: 83800ACD66B2B9 CRC64:			



DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.  
DR PROSITE: PS00004; C2\_DOMAIN\_2; 2.  
SQ SEQUENCE 403 AA; 45482 MW; 3153ED7EC1DEE9FB CRC64;

Query Match 28.5%; Score 627.5; DB 11; Length 403;  
Best Local Similarity 36.2%; Pred. No. 3,76-38;  
Matches 133; Conservative 72; Mismatches 127; Indels 71; Gaps 11;

QY 27 VFTVSL-----FAMIC--CORSSKSNKTPPYKFFVHLKGVDIYPENLSKKKEGAD-- 77  
DB 23 IITVSLSVITVLCGLCHWCORLGRRY-----NSLETVGTPDSG 62  
QY 78 -----KNEVK-----NKPAVPKNSLHLDLEKRDINGNFPKTNLKGSPSDLENATPK 124  
DB 63 RGRGEKKAIRLPAGKAVNTAPVPGQTPHDSDDR-----TEPRS-----SVSDLVN---- 109  
QY 125 LFLGEEKSVSPESLKSSTSLTSEB-----KQEKLGTLFFSLFLENERKAFVNIKEARG 179  
DB 110 -----SLTSEMLTLSPGSEEDAHGCSRENIGRIOSVGNFOESTLTVKVKAOE 161  
QY 180 LPAMDQSMSTDPYIKMTILPEKKHKVTRVLRKTLDPAFDETFEYGIPTQIOELALH 239  
DB 162 LPKAD-FSGSDPEFVKIYLLPDKKHLETYKVRKNLNPNNETFLFEGFPEYKVVORILY 220  
QY 240 FTLSFDRFSRDIIGEVLPILSGIELSEGMKLNREIRKRVKSSSGRGELLISLCYOS 299  
DB 221 LOYLDVDFRSNDPIGEVSIPLNKVDLTQ--MOTFMKDLKPCSDSGSGRELLISLCYNP 278  
QY 300 TTNLTLYVYVKAHRLPKSDVSGISDPYKVNLYHAKKRISKKTTHVKKCTPNAVFNELEFV 359  
DB 279 SANSIIVNITKARNLKAAMDIGSTDPYKVMYKDKRVEKKKTVKRRNLINFESPA 338  
QY 360 FDPICELEIDIEVEFLVDSESGSNREVIYIGOLVGAAGEGGCHMEICDYPRIQIAW 419  
DB 339 FDIPTELKRETTIITVMDKRLSRNDVIYIGKITLWSKSGGEYKHNKMDIARPRQVAAQM 398  
QY 420 HVL 422  
DB 399 HOL 401

RESULT 10  
ID 025393 PRELIMINARY; PRT: 424 AA.  
AC 025393;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DE SYNACTOTAGMIN.  
OS Loligo pealeii (Longfin squid).  
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Teuthoidea;  
OX NCBI\_TaxId=6621;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=OPTIC LOBE;  
RX MEDLINE=9606884; PubMed=7479868;  
RA Mikoshiba K., Fukuda M., Moreira J.E., Lewis F.M.T., Sugimori M.,  
RA Nlinobe M., Llinas R.;  
RT "Role of the C2A domain of synaptotagmin in transmitter release as  
RT determined by specific antibody injection into the squid giant synapse  
RT preterminal."  
RL Proc. Natl. Acad. Sci. U.S.A. 92:10703-10707(1995).  
DR EMBL: D63797; RASV. 409866.1; -  
DR HSSP: P21707; IRASV.  
DR InterPro: IPR000008; C2.  
DR InterPro: IPR002149; LRI.  
DR Pfam: PF00168; C2; 2.  
DR PRINTS: PR00360; C2DOMAIN.  
DR PRINTS: PR00399; SYNACTOTAGMIN.  
DR SMART: SM00239; C2; 2.

DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.  
DR PROSITE: PS00004; C2\_DOMAIN\_2; 2.  
SQ SEQUENCE 424 AA; 47056 MW; F9733D3B10CF93C0 CRC64;

Query Match 28.5%; Score 627; DB 5; Length 424;  
Best Local Similarity 34.6%; Pred. No. 4,36-38;  
Matches 144; Conservative 75; Mismatches 137; Indels 60; Gaps 8;

QY 10 EFDEIPTVVGIFSAAGLFTVSLFPAWICQKSSKSNKTPPYKFFVHLKGVDIYPENLNS 69  
DB 59 ELEKPIPAITILICAGVILFLVCGTYCCCKRICRRRG----- 96  
QY 70 KKKFGADKNEVKNKPAVPKNSLHLDLEKRDINGNFPKTNLKGSPSDLENATPKLELEG 129  
DB 97 -----KDKRKGKGA-----VDLKGVLGNSIKERKVP--DLEEL----- 130  
QY 130 EKESVSPESLKSSTSLTSEKQEKLTGTLFFSLFLENERKAFVNIKEARGLPAMDQSM 189  
DB 131 -----PNMEDNEDAEASTSEVYLGKIQYSDYDPOKGLTVNVIQAADPQMD-MSGT 183  
QY 190 SDPYIKMTILPEKKHKVTRVLRKTLDPAFDETFEYGIPTQIOELALHFTLSFDRFS 249  
DB 184 SDPYKVVYVLMPPDKKKKFKETKVRKTLNPVNESFTPKNVPAADITGKTLVPAIYDFDRFS 243  
QY 250 RDIIGEVLPILSGIELSEGMKLN-REIRK--RNVKSSSGRELLISLCYOSTTNTLT 306  
DB 244 KHDQIGOVVAMNSIDL--GSVMEWRDLTSPDNDKRNKLGICFSLRVVPAKGLTV 301  
QY 307 VYLRARHLPKSDVSGISDPYKVNLYHAKKRISKKTTHVKKCTPNAVFNELEFVIDPEG 366  
DB 302 VILEAKNLKMDVGLSDPYKISLMLNGRIKKKTKTYKCTLNPYNESFAFEVPEEQ 361  
QY 367 LDDIVVEFLVDSESGSNREVIYIGOLVGAAGEGGCHMEICDYPRIQIAW 422  
DB 362 IKGSLVYTVVDYDRIGTSEPTIGTFTGCNSTGTGLRHNSDMLNRRPRVAAQMTL 417

RESULT 11  
ID 09VOC7 PRELIMINARY; PRT: 474 AA.  
AC 09VOC7;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE SYT PROTEIN.  
GN SYT OR G3139.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxId=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame C.R., Miklos G.L.G.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Andrews-Pfannkoch C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Betman B.P., Bhandarai D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavalier S., Dahlke C., Davenport I.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doull L.E., Downes M., Dugan-Rocha S., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan F., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibeagwa C.,  
 RA Jatali D., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003582; AAF51205.1; -.  
 DR HSSP: P21707; IBYN.  
 DR FLYBASE: FBgn0004242; syt.  
 DR InterPro: IPR000008; C2.  
 DR InterPro: IPR002149; LRI.  
 DR InterPro: IPR001565; Synaptotagmin.  
 DR Pfam: PF00168; C2; 2.  
 DR PRINTS: PR00360; C2DOMAIN.  
 DR PRINTS: PR00399; SYNAPTOTAGMN.  
 DR SMART: SM00239; C2; 2.  
 DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.  
 DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.  
 DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.  
 SQ SEQUENCE 474 AA; 53260 MW; 76F3A34EABE875B CRC64;

Query Match 28.3%; Score 622.5; DB 5; Length 474;  
 Best Local Similarity 34.7%; Pred. No. 1,1e-37;  
 Matches 151; Conservative 67; Mismatches 144; Indels 73; Gaps 9;

3 PITTSREDFEPIYV-----GIFSAGLVETVSLFAMICCRKSSKSNKTPPK 52  
 88 PVIKRIEHGVETVEVIAERTGPTGVAIIIVLVFGIIFVRRFLKKRRK--- 144  
 53 FVHVLKGVDTIYENLNKSKKFGADKNEVKNKPAVPKNSLHDLERKDLNGNPPKTNLKP 112  
 145 -----DGKKKG-----VDKSVOLLGSAVKKKQVP 170  
 113 GSPSDLENATPKLFEGEKESVPSLSKSTSLTSEKOEKLGTLFSLSEYNERKAFV 172  
 171 DMBELTENAE-----EGDE-----DKOSECKLGRLFKLETPDNSSLAV 211  
 173 NIKKARGPAMDQSMSTSDPIKMTILPEKKHKVKTIVLRKTLDPADFETFTYGIPTQ 232  
 212 TVIQAEELPALD-MGQTSPPYKVVLLPDKKKKFKVHRKTLSPVFNFTFFKSLPYAD 270  
 233 IOBLAHFTILSFDRSRDITIGEVILPLSGIELSEGMKLMNRITKRVKSSGR---G 289  
 271 AMNKTLVFAIFEDRSPKSDQIGEVKVPCLCTIDLQ-TIEEMRDV--SVEEGSGGEKIG 327  
 290 ELLISLCYOSTNTLTVVVYLKARHLKPSDVSGLSDPYVKNLYHAKKRISKKTHYKKT 349  
 328 DICFSRIYPTAGKLTIVLLEAKNLKMDVGGISDPYVAIAIMONKRLKKKTKSTKCT 387  
 350 PNAVNEFLVFDPICBGLDISEVFLVDSERGSRENVIGQVLVGAAGGEGEHKKEIC 409  
 388 LNPYNESEFVEPFOIOKICLVVVVDYDRIGTSEPIGRCLIGCMGCTELRHMSDML 447  
 410 DYPRROIATKWHVCD 424  
 448 ASFRPPIAQMHTLKD 462

RESULT 12  
 007529 PRELIMINARY; PRT: 403 AA.  
 AC 007529  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE SYNAPTOTAGMIN (FRAGMENT).  
 OS Loligo pealeii (longfin squid).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Teuthoidea;  
 OC Myopsida; Loliginidae; Loligo.  
 OX NCBI\_Taxid=6621;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-93247639; PubMed-8097867;  
 RA Bommer K., Charlton M.P., De Bello W., Chin G., Betz H.,  
 RA Augustine G.J.;  
 RT "Inhibition of neurotransmitter release by C2-domain peptides  
 RT implicates synaptotagmin in exocytosis.";  
 RL Nature 363:163-165(1993).  
 CC -I- FUNCTION: MAY HAVE A REGULATORY ROLE IN THE MEMBRANE INTERACTIONS  
 CC DURING TRAFFICKING OF SYNAPTIC VESICLES AT THE ACTIVE ZONE OF THE  
 CC SYNAPSE. IT BINDS ACIDIC PHOSPHOLIPIDS WITH A SPECIFICITY THAT  
 CC REQUIRES THE PRESENCE OF BOTH AN ACIDIC HEAD GROUP AND A DIACYL  
 CC BACKBONE.  
 CC -I- SUBUNIT:: MONOMER (PROBABLE).  
 CC -I- SUBCELLULAR LOCATION: SYNAPTIC VESICLES AND CHROMAFFIN GRANULES.  
 DR EMBL: X72386; CAA51079.1; -.  
 DR HSSP: P21707; IBYN.  
 DR InterPro: IPR000008; C2.  
 DR InterPro: IPR002149; LRI.  
 DR InterPro: IPR001565; Synaptotagmin.  
 DR Pfam: PF00168; C2; 2.  
 DR PRINTS: PR00360; C2DOMAIN.  
 DR PRINTS: PR00399; SYNAPTOTAGMN.  
 DR SMART: SM00239; C2; 2.  
 DR PROSITE: PS00499; C2\_DOMAIN\_1; 2.  
 DR PROSITE: PS50004; C2\_DOMAIN\_2; 2.  
 KW Transmembrane; Synapse.  
 FT NON\_TER 1  
 FT TRANSMEM 46  
 FT SEQUENCE 403 AA; 45434 MW; 74ADF87A0B5FD838 CRC64;

Query Match 28.2%; Score 621; DB 5; Length 403;  
 Best Local Similarity 35.6%; Pred. No. 1,1e-37;  
 Matches 149; Conservative 71; Mismatches 131; Indels 68; Gaps 10;

10 EDEIDPTVVGIFSAGLVETVSLFAMICQ---RKSSKSNKTPPKFVHVLKGVDTIYPEN 66  
 40 ELEKPIWAILICAGVLLFVCGTCCCKRICRRGKDGKGLGAVDLRGVOLL--- 96  
 67 LNSKKKFGADKNEVKNKPAVPKNSLHDLERKDLNGNPPKTNLKGSPSDLENATPKLF 126  
 97 -----GNSIKKP-----DLEELPMN-----MEDNEDA----- 119  
 127 LSEKESVPSLSKSTSLTSEKOEKLGTLFSLSEYNERKAFVKNIKKARGPAMDQ 186  
 120 -----ESTKS-----EVLKGLKIOYSMDYDFQKGLGVNVAADLPAMD-M 159  
 187 SMSTSDPIYKMTILPEKKHKVKTIVLRKTLDPADFETFTYGIPTQIOELAHFTLSPD 246  
 160 SGTSDPYVAVVYVLPMDKKKFKETVHKRKTLPVFNESFTFRNVPYADITGKTLVFAIYDD 219  
 247 RESRDDIIEVILPLSGIELSEGMKLMN-REIK--RNVRKSSRGELLSLCYSTTWN 303  
 220 RESKHQIOGOVYAMNSIDL--GSVMEWRDLTSPDDAKEKKKLTGICFSLRYPVTAK 277  
 304 LTVVYLKARHLKPSDVSGLSDPYVKNLYHAKKRISKKTHYAKKCPNPAVNEFLVFDLP 363  
 278 LTVVILKARHLKMDVGGISDPYKISLMLNGRKIRKKTKTVAKCLNPNYNESFAFEPV 337

OY 364 CEGLEDISEVFLVDSERGSNEVIGOLVGAAGTGEHKEICDYPRIQIAKMHVL 422  
 Db 338 FEQIQKVSILVTVVVDYDRHMTSEPIGRFLGCSNIGTGLRHMSMLANPRPIAQNHTL 396

## RESULT 13

AC 043581 PRELIMINARY; PRT: 418 AA.

ID 043581  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE SYNAPTOTAGMIN VII (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cooper P.R., Nowak N.J., Higgins M.J., Church D.M., Shows T.B.;  
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF038535; AAB92667.1;  
 DR HSSP; P21707; IRVY.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR002149; LRI.  
 DR Pfam; PF00168; C2.2.  
 DR PRINTS; PRO0360; C2DOMAIN.  
 DR SMART; SM00339; SYNAPTOTAGMIN.  
 DR PROSITE; PS00499; C2\_DOMAIN\_1; 2.  
 DR PROSITE; PS50004; C2\_DOMAIN\_2; 2.  
 FT NON TER 1  
 SO SEQUENCE 418 AA; 47176 MW; 576509854E2B5FF CRC64;

Query Match 28.0%; Score 615.5; DB 4; Length 418;  
 Best Local Similarity 36.2%; Pred. No. 3e-37;  
 Matches 153; Conservative 69; Mismatches 130; Indels 71; Gaps 11;

OY 27 VETVSL-----FAMIC--CORRKSNNKTPRYKVVHLKGVDIYENLNKKKFGAD-- 77  
 Db 38 IITVSLVTVVVLGLCHMCORIKGRKK-----NSLETGVGPDG 77  
 OY 78 -----KNEVK-----NKPAVKNSLHLEKRDNLGNFPKTNLKPSPDLENATPK 124  
 Db 78 RGRSEKAKILPAGKAVMTAPVPGQTPHDESDR--TEPRS-----SVSDLVN---- 124  
 OY 125 LFLGKEKESVPESLSKSTSLTSEE-----KQKLGTLFSLFLEYNFPRKAFVYNKEARG 179  
 Db 125 -----SLTSEMLSLSPGSEDEAHGCSRENIGRQFVGINPQESLTLTKIAKAOE 176  
 OY 180 LPAMDQSMTPSDPYIKMTILPEKKHKVTRVLRKTLDPADPETFYFGITPTIOELAH 239  
 Db 177 LPAMD-FSGTSDPFVKIYLPDKKHKVTRVLRKTLDPADPETFYFGITPTIOELAH 235  
 OY 240 FTILSPDRSRDIIIGEVLLPLSGIELSEGMMLNREIKRNVKSSGSGELLISLCYOS 299  
 Db 236 LOYLDYDRSRNDPIGEVSLPLNKVDLTO--MOTFWKDLKPCSGSGSGELLISLCYNP 293  
 OY 300 TTTTLTVVVLKARHLPRKSDVSGLSDPYKVLNLAHAKRISKTKTHVKKCTPNAVNELEFV 359  
 Db 294 SANSIIVNIIRKARNLKAMDIGTSDPYKVLNLAHAKRISKTKTHVKKCTPNAVNELEFV 353  
 OY 360 FDIPEGLEDISEVFLVDSERGSNEVIGOLVGAAGTGEHKEICDYPRIQIAKMHVL 419  
 Db 354 FDIPEKLETRITIIITVMDKILSRNDVIKIIYLSKMSGGEVYKHKMDIARPRQVPAQM 413  
 OY 420 HVL 422  
 Db 414 HOL 416

## RESULT 14

ID 099P37 PRELIMINARY; PRT: 520 AA.

AC 099P37  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE SYNAPTOTAGMIN VIIA.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sugita S., Han W., Butz S., Fernandez-Chacon R., Lao Y., Sudhof T.C.;  
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF336856; AAK01451.1;  
 SO SEQUENCE 520 AA; 57441 MW; 409B9696CD46B502 CRC64;

Query Match 28.0%; Score 615.5; DB 11; Length 520;  
 Best Local Similarity 38.6%; Pred. No. 4e-37;  
 Matches 140; Conservative 70; Mismatches 122; Indels 31; Gaps 7;

OY 65 ENLNKKKFGADKKNEVKNKPAVKNSLHLEKRDNLGNFPKTNLKPSPDLENATPK 124  
 Db 182 QNAGGPKRLPAGK--AVNTAPVPGQTPHDESDR--TEPRS-----SVSDLVN---- 226  
 OY 125 LFLGKEKESVPESLSKSTSLTSEE-----KQKLGTLFSLFLEYNFPRKAFVYNKEARG 179  
 Db 227 -----SLTSEMLSLSPGSEDEAHGCSRENIGRQFVGINPQESLTLTKIAKAOE 278  
 OY 180 LPAMDQSMTPSDPYIKMTILPEKKHKVTRVLRKTLDPADPETFYFGITPTIOELAH 239  
 Db 279 LPAMD-FSGTSDPFVKIYLPDKKHKVTRVLRKTLDPADPETFYFGITPTIOELAH 237  
 OY 240 FTILSPDRSRDIIIGEVLLPLSGIELSEGMMLNREIKRNVKSSGSGELLISLCYOS 299  
 Db 338 LOYLDYDRSRNDPIGEVSLPLNKVDLTO--MOTFWKDLKPCSGSGSGELLISLCYNP 395  
 OY 300 TTTTLTVVVLKARHLPRKSDVSGLSDPYKVLNLAHAKRISKTKTHVKKCTPNAVNELEFV 359  
 Db 396 SANSIIVNIIRKARNLKAMDIGTSDPYKVLNLAHAKRISKTKTHVKKCTPNAVNELEFV 455  
 OY 360 FDIPEGLEDISEVFLVDSERGSNEVIGOLVGAAGTGEHKEICDYPRIQIAKMHVL 419  
 Db 456 FDIPEKLETRITIIITVMDKILSRNDVIKIIYLSKMSGGEVYKHKMDIARPRQVPAQM 515  
 OY 420 HVL 422  
 Db 516 HOL 518

## RESULT 15

ID 099P36 PRELIMINARY; PRT: 523 AA.

AC 099P36  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE SYNAPTOTAGMIN VIIA.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sugita S., Han W., Butz S., Fernandez-Chacon R., Lao Y., Sudhof T.C.;  
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 12, 2001, 10:53:49 ; Search time 14.4 Seconds  
(without alignments)  
664.160 Million cell updates/sec

Title: US-09-680-121-2

Perfect score: 2201

Sequence: 1 MADITTSREBFDEIPYVGI.....KEICDYPRIQIAKHWLDCG 425

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/6C.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2201	100.0	425	US-09-036-315-2	Sequence 2, Appl
2	2006	91.1	425	US-09-036-315-5	Sequence 5, Appl
3	605	27.5	121	US-09-036-315-7	Sequence 7, Appl
4	581	26.4	422	US-08-872-979-8	Sequence 8, Appl
5	577	26.2	113	US-09-036-315-10	Sequence 10, Appl
6	361.5	16.4	355	US-08-872-979-7	Sequence 7, Appl
7	338.5	15.4	375	US-08-872-979-3	Sequence 3, Appl
8	310	14.1	60	US-09-036-315-24	Sequence 24, Appl
9	288	13.1	138	US-08-609-049A-17	Sequence 17, Appl
10	288	13.1	138	US-09-170-996-17	Sequence 17, Appl
11	271	12.3	120	US-09-036-315-9	Sequence 9, Appl
12	244	11.1	115	US-09-036-315-8	Sequence 8, Appl
13	226	10.3	46	US-09-036-315-25	Sequence 25, Appl
14	182.5	8.3	140	US-08-609-049A-16	Sequence 16, Appl
15	182.5	8.3	140	US-09-170-996-16	Sequence 16, Appl
16	180	8.2	34	US-09-036-315-22	Sequence 22, Appl
17	169	7.7	33	US-09-036-315-26	Sequence 26, Appl
18	166	7.5	31	US-09-036-315-21	Sequence 21, Appl
19	158	7.2	1658	US-08-609-049A-13	Sequence 13, Appl
20	158	7.2	1658	US-09-170-996-13	Sequence 13, Appl
21	158	7.2	1726	US-08-609-049A-30	Sequence 30, Appl
22	158	7.2	1726	US-09-170-996-30	Sequence 30, Appl
23	157.5	7.2	136	US-08-609-049A-18	Sequence 18, Appl
24	157.5	7.2	136	US-09-170-996-18	Sequence 18, Appl
25	155.5	7.1	117	US-09-036-315-6	Sequence 6, Appl
26	152.5	6.9	1876	US-08-609-049A-12	Sequence 12, Appl
27	152.5	6.9	1876	US-09-170-996-12	Sequence 12, Appl

28	149.5	6.8	137	2	US-08-609-049A-14	Sequence 14, Appl
29	149.5	6.8	137	2	US-09-170-996-14	Sequence 14, Appl
30	144.5	6.6	1876	4	US-08-609-049A-28	Sequence 28, Appl
31	144.5	6.6	1876	4	US-09-170-996-28	Sequence 28, Appl
32	141	6.4	671	6	5266464-2	Patent No. 5266464
33	136	6.2	804	3	US-08-909-954-2	Sequence 2, Appl
34	131.5	6.0	799	3	US-08-909-954-4	Sequence 4, Appl
35	129.5	5.9	137	2	US-08-609-049A-15	Sequence 15, Appl
36	129.5	5.9	137	4	US-09-170-996-15	Sequence 15, Appl
37	125	5.7	137	4	US-09-036-315-23	Sequence 23, Appl
38	117	5.3	927	3	US-08-895-601-6	Sequence 6, Appl
39	109	5.0	990	2	US-08-392-625-20	Sequence 20, Appl
40	109	5.0	990	2	US-08-466-961A-20	Sequence 20, Appl
41	102.5	4.7	990	2	US-08-645-193B-15	Sequence 15, Appl
42	97.5	4.4	290	2	US-08-576-626A-34	Sequence 34, Appl
43	97	4.4	747	3	US-09-035-648-18	Sequence 18, Appl
44	97	4.4	747	4	US-09-001-951-18	Sequence 18, Appl
45	96	4.4	506	2	US-08-929-501-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-09-036-315-2  
; Sequence 2, Application US/09036315  
; Patent No. 6218523  
; GENERAL INFORMATION:  
; APPLICANT: French, Cynthia K.  
; APPLICANT: Schneider, Patrick A.  
; APPLICANT: Yamamoto, Karen K.  
; TITLE OF INVENTION: Prostate Cancer-Specific Marker  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/036,315  
; FILING DATE: 06-MAR-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/047,811  
; FILING DATE: 15-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/041,246  
; FILING DATE: 07-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Storella, John R.  
; REGISTRATION NUMBER: 32,944  
; REFERENCE/DOCKET NUMBER: 018002-00021005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 425 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-036-315-2  
Query Match 100.0%; Score 2201; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 4.4e-217;

Fri Dec 14 10:32:11 2001

us-09-680-121-2.rai

Page 2

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MADITTSREDFEIPVVGISAFGLVFWYSLFAMITCCKRSSKSNKPPYKFYVHLGV	60
Db	1	MADITTSREDFEIPVVGISAFGLVFWYSLFAMITCCKRSSKSNKPPYKFYVHLGV	60
Qy	61	DIYPENINSKKRFGADDKKVEYNKPAVPKNSLHJDEKRLNMFNPKTJLKRGSSDLEN	120
Db	61	DIYPENINSKKRFGADDKKVEYNKPAVPKNSLHJDEKRLNMFNPKTJLKRGSSDLEN	120
Qy	121	ATPKLFLEGGKESVSPESHKSTSTLTSSEKOEKGLTFESLEYNEPKAFVYNIKEARGL	180
Db	121	ATPKLFLEGGKESVSPESHKSTSTLTSSEKOEKGLTFESLEYNEPKAFVYNIKEARGL	180
Qy	181	PAMDQSMSTDYIKMTLIPKKKKHVTRVLRKTRDPAPDEFPTFGIPRYOIGLALHF	240
Db	181	PAMDQSMSTDYIKMTLIPKKKKHVTRVLRKTRDPAPDEFPTFGIPRYOIGLALHF	240
Qy	241	TITSPRBSRDITIGEVLPISGLTISBGMKMLNRELIRKNVAKSSGRELILISCTOYST	300
Db	241	TITSPRBSRDITIGEVLPISGLTISBGMKMLNRELIRKNVAKSSGRELILISCTOYST	300
Qy	301	TNTLVYVYLKAARLKRSDVSGSLSDPYRYKVNLYHAKKRISKSKTKTHVKKCTPNAVFNELFV	360
Db	301	TNTLVYVYLKAARLKRSDVSGSLSDPYRYKVNLYHAKKRISKSKTKTHVKKCTPNAVFNELFV	360
Qy	361	DIFCEGLIEDISVEFLVLDSERGSRNEVIGOLVGAAAEGTGEHMKETICDYPRRIOAKWH	420
Db	361	DIFCEGLIEDISVEFLVLDSERGSRNEVIGOLVGAAAEGTGEHMKETICDYPRRIOAKWH	420
Qy	421	VLCDG 425	
Db	421	VLCDG 425	

RESULT 2  
US-09-036-315-5  
Sequence 5, Application US/09036315  
Patent No. 6218523  
GENERAL INFORMATION:  
APPLICANT: French, Cynthia K.  
APPLICANT: Schneider, Patrick A.  
APPLICANT: Yamamoto, Karen K.  
TITLE OF INVENTION: Prostate Cancer-Specific Marker  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/036,315  
FILING DATE: 06-MAR-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/047,811  
FILING DATE: 15-MAY-1997  
PRIOR APPLICATION NUMBER: US 60/041,246  
APPLICATION DATE: 07-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 018002-00021005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200

```

? TELEFAX: (415) 576-0300
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 425 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? FEATURE:
? NAME/KEY: Protein
? LOCATION: 1..425
? OTHER INFORMATION:
? OS-09-036-315-5 /note="rat synaptotagmin 4 (SYT4)"

```

Query Match	91.1%;	Score 2006;	DB 4;	Length 425;
Best Local Similarity	89.9%;	Pred. No. 4.2e-197;		
Matches 382;	Conservative 21;	Mismatches 22;	Indels 0;	Gaps 0

QY	1	MAPITTSSEEDIEIPVYVIGIFASAGLWFTVLSLPAWICCORSSKSNTPPYKRVHVLGV	60
Db	1	MAPITTSVEEDEIPVYVIGIFASGLVFTVLSLPAWICCORSAKSNTPPYKRVHVLGV	60
QY	61	DYEPENLUSKKKFGADDKREYKNRPVAPKNSLHLDKERDNLGNFPKTNLKPSPSDLEN	120
Db	61	DYEPENLUSKKKFGGDKSEAKRAALPNLSLHLDKERDNLGNFPKTNPKASSSDLEN	120
QY	121	APPKFLFEGKESEVSESLKSTSLTSEBQOKLGTLPSELEYNEPKAPVYVYKAEGL	180
Db	121	VYPKFLPETERKEAVSESLKSTSLTSEBQOKLGTLPSELEYNEPKAPVYVYKAEGL	180
QY	181	PAMBOQMSSTDPIKNTILPEKKHKVTVRLKRTLDPADEFYTFYCIPTQLOELAHF	240
Db	181	PAMBOQMSSTDPIKNTILPEKKHKVTVRLKRTLDPADEFYTFYCIPTQLOELSLHF	240
QY	241	TILSFDRFSRDIIGVLLPLSGIBLSEGMMLNRETIKRVKRSRGEGLLSLSLCYOST	300
Db	241	TVLSDRFSRDIIGVLLPLSGIBLSDGMMLTRETIKRVKRSRGEGLLSLSLCYOST	300
QY	301	TNTLTLYVYLKAARLPKSDVSGSLSDPYKVALYAKKRRISKKTKYVKKCPNAVENDELVF	360
Db	301	TNTLTLYVYLKAARLPPSDVSGSLSDPYKVALYAKKRRISKKTKYVKKCPNAVENDELVF	360
QY	361	DIPCEGLIEDISVEFLVLSERGSRNEVIGQLVLGAAGCTGGEHMKEICDYPPRROIAWMH	420
Db	361	DIPCEGLIEDISVEFLVLSERGSRNEVIGRLVLGATAGSGGHHKEICDYPPRROIAWMH	420
QY	421	VLCDG 425	
Db	421	MLCDG 425	

RESULT 3  
US-09-036-315-7  
Sequence 7, Application US/09036315  
Patent No. 6218523  
GENERAL INFORMATION:  
APPLICANT: French, Cynthia K.  
APPLICANT: Schneider, Patrick A.  
APPLICANT: Yamamoto, Karen K.  
TITLE OF INVENTION: Prostate Cancer-Specific Marker  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
INVENTOR: Thompson and Townsend and Crew LLP  
FIRM: Thompson and Townsend and Crew LLP  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,315
FILING DATE: 06-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,811
FILING DATE: 15-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,246
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 018002-000210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 121 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..121
OTHER INFORMATION: /note="Repro-PC-1.0 (PC-20) "B"
OTHER INFORMATION: Internal repeat (amino acid
OTHER INFORMATION: positions 276-397)"
US-09-036-315-7
```

Query Match 27.5%; Score 605; DB 4; Length 121;

Best Local Similarity 99.2%; Pred. No. 2e-54; Mismatches 1; Indels 0; Gaps 0;

```
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 276 EIKRNRVRSKSGGELLISLCYOSTNTLTJVVLKARHLPKSDVSGLSDPYKVNLYHAK 335
DB 1 EIKRNRVRSKSGGELLISLCYOSTNTLTJVVLKARHLPKSDVSGLSDPYKVNLYHAK 60
QY 336 KRISKKTHVKKCTPNAVFNELFVDPICEGLDISVEFLVLDSEKSRNEVIGQLVLA 395
DB 61 KRISKKTHVKKCTPNAVFNELFVDPICEGLDISVEFLVLDSEKSRNEVIGQLVLA 120
QY 396 A 396
DB 121 A 121
```

```
RESULT 4
US-08-872-979-8
Sequence 8, Application US/08872979
Patent No. 6074844
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,979
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0320 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 422 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 338658
US-08-872-979-8
```

Query Match 26.4%; Score 581; DB 3; Length 422;

Best Local Similarity 39.9%; Pred. No. 4.3e-51; Mismatches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;

```
Matches 132; Conservative 65; Mismatches 114; Indels 20; Gaps 8;
QY 105 FPKTNLKPSPSPDLNATPKLFEGEKESVPSK---STSLT-SEK-----OEKIG 155
DB 85 FKKNNKKKKEKGGKNAIMNKDVKDLGKTKKDALDDDAEFTLTGEEKKEEKKEKIG 144
QY 156 TLEFSLEYNFERRAFVNIKEARGLPAMDQSMTPYIKMTLIPKKKHVKTRVLRKTL 215
DB 145 KLDYSLDYDFQNNQLLVGIQIAELPALD-MGTSDPYVVFLLPPKKKFEKVRHRTL 203
QY 216 DPAFDETFYFGIPYQIQELAHFTILSPDRSRDIDIGEVLIPLSGIEL-----SEGM 271
DB 204 NPVFNQFTF-KVYSSELGKTLVMAVYDFDRSKHDIIGEFVPMNVYDFGHVTEWRD 262
QY 272 LMRREIKRVRKSSGGELLISLCYOSTNTLTJVVLKARHLPKSDVSGLSDPYKVN 331
DB 263 LQSAE--KEQEK---LGDICFSLRYVPTAGKLTIVYLEKKNKKMDVSGLSDPYKIH 317
QY 332 YHAKKRISKKTHVKCTPNAVFNELFVDPICEGLDISVEFLVLDSEKSRNEVIGOL 391
DB 318 MONGKRLLKKKTKTKNTLNPYVNESFVPEQIQKVQVVTVLDPYKIGKDAIGKV 377
QY 392 VLGAAAGTGGGEHWKELCDYPRQIAKMYL 422
DB 378 FVGYNSTGAELRHMSDMLNRRPPIQWHTL 408
```

```
RESULT 5
US-09-036-315-10
Sequence 10, Application US/09036315
Patent No. 6218523
GENERAL INFORMATION:
APPLICANT: French, Cynthia K.
APPLICANT: Schneider, Patrick A.
APPLICANT: Yamamoto, Karen K.
TITLE OF INVENTION: Prostate Cancer-Specific Marker
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk.
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,315
FILING DATE: 06-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,811
FILING DATE: 15-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,246
FILING DATE: 07-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 018002-000210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..113
OTHER INFORMATION: /note="Repro-PC-1.0 (PC-20) "A"
OTHER INFORMATION: Internal repeat (amino acid
positions 150-263)"
US-09-036-315-10

Query Match 26.2%; Score 577; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.3e-51;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 KOEKLGLFSLSEYENERRKAFVYNIKEARGLPAMDEOSMTSDPYIKMTILPEKKHKVKT 209
DB 1 KOEKLGLFSLSEYENERRKAFVYNIKEARGLPAMDEOSMTSDPYIKMTILPEKKHKVKT 60
QY 210 VLKRTLPARDEFTEFYGIPTQIOELALHFTILSFDRFSRDDIIGEVLLPLS 262
DB 61 VLKRTLPARDEFTEFYGIPTQIOELALHFTILSFDRFSRDDIIGEVLLPLS 113

RESULT 6
US-08-872-979-7
Sequence 7, Application US/08872979
Patent No. 6074844
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,979
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0320 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 643658
US-08-872-979-7

Query Match 16.4%; Score 361.5; DB 3; Length 355;
Best Local Similarity 30.2%; Pred. No. 1e-28;
Matches 94; Conservative 62; Mismatches 120; Indels 35; Gaps 6;

QY 130 EKESYRESLKSTSLTSEER-----QEKIGTLFSLSEYENERRKAFVYNIKEAR 178
DB 39 DKETVLGSARSTHTLVDPDVCLEPCSGGDOMKRLSLIEDFGSETRVLKATL 98
QY 179 GLPAMDEOSMTSDPYIKMTILPEKKHKVKTIVLAKTLPARDEFTEFYGIPTQIOELAL 238
DB 99 NLKA---EGTADPYAMVSVTSGGRHETKVNHGTLSPWETCCFL-VPPAELPKATL 153
QY 239 HTTILSFDRFSRDDIIGEVLLPLSGIELSECKMLMREITKR-----NVKSSRGEL 291
DB 154 KVALMDFKRFSEHEPLGELPLDGTVDL-----QHVLESWYOLGPPGTPEPOMGEL 205
QY 292 LLSICVOSTTNTLVVVKARHLKPSDVSGLSDPYKVNLYNAKKRISKKTHTVKKCTPN 351
DB 206 CFSRLRVPSGSLTVVLEARGLN---PGLAAYVAKIOLMLNORKKSKTSKKTGTT 261
QY 352 AVFNELEFVDCIEGLIEDISVEFLVLDSEGRSNEVIGQVLGAAGTGGEHMEICDY 411
DB 262 PYENAEFVFLVVSOLQSDVLDVLAVMARGQLRTEPVGKVLGSRASGROPLOHWMADLH 321
QY 412 PRQIAKMHVL 422
DB 322 ARRPDAQMHHL 332

RESULT 7
US-08-872-979-3
Sequence 3, Application US/08872979
Patent No. 6074844
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

```

RESULT 8  
US-09-036-315-24  
Sequence 24, Application US/09036315  
Patent No. 6218523  
GENERAL INFORMATION:  
APPLICANT: French, Cynthia K.  
APPLICANT: Schneider, Patrick A.  
APPLICANT: Yamamoto, Karen K.  
TITLE OF INVENTION: Prostate Cancer-Specific Marker  
NUMBER OF SEQUENCES: 27  
CORRESPONDING ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP

RESULT 9  
 ; US-08-609-049A-17  
 ; Sequence 17, Application US/08609049A  
 ; Patent No. 5948664  
 ;  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Williams, Lewis T.  
 ; APPLICANT: Molz, Lisa  
 ; APPLICANT: Chen, Yen-Men  
 ; TITLE OF INVENTION: No. 5948664e1 PI 3'-kinase Polypeptides  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ;  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ;  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/609,049A  
 ; FILING DATE: 29-FEB-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:

NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-063700US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-609-049A-17

Query Match 13.1%; Score 288; DB 2; Length 138;  
Best Local Similarity 46.8%; Pred. No. 7,6e-22;  
Matches 59; Conservative 22; Mismatches 45; Indels 0; Gaps 0;

OY 289 GELLISLCYOSTNTITVYVVKARHLPRKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKC 348  
DB 12 GDICTSLRVPYPTAGKLVLCILEAKNLMKMDVGLSDPYKIHLMONGKRLLKKTTHVKKK 71  
OY 349 TPNAVFNELFVFDICPEGLEDISVEFLVLDSEGRSRENYIGQLVGAAGCTGGEHMKKEI 408  
DB 72 TLNPTNSESFEIPEFDIQKVQVYVTVLDYDKLGKNEALIKIFVGSNATGTLELRHMSDM 131  
OY 409 CDYPRR 414  
DB 132 LANPRR 137

RESULT 10  
US-09-170-996-17  
Sequence 17, Application US/09170996  
Patent No. 6291220  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Moitz, Lisa  
TITLE OF INVENTION: Chen, Yen-Wen  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/170,996  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/609,049  
FILING DATE: 29-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-063700US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-170-996-17

Query Match 13.1%; Score 288; DB 4; Length 138;  
Best Local Similarity 46.8%; Pred. No. 7,6e-22;  
Matches 59; Conservative 22; Mismatches 45; Indels 0; Gaps 0;

OY 289 GELLISLCYOSTNTITVYVVKARHLPRKSDVSGLSDPYKVNLYHAKKRISKKTTHVKKC 348  
DB 12 GDICTSLRVPYPTAGKLVLCILEAKNLMKMDVGLSDPYKIHLMONGKRLLKKTTHVKKK 71  
OY 349 TPNAVFNELFVFDICPEGLEDISVEFLVLDSEGRSRENYIGQLVGAAGCTGGEHMKKEI 408  
DB 72 TLNPTNSESFEIPEFDIQKVQVYVTVLDYDKLGKNEALIKIFVGSNATGTLELRHMSDM 131  
OY 409 CDYPRR 414  
DB 132 LANPRR 137

RESULT 11  
US-09-036-315-9  
Sequence 9, Application US/09036315  
Patent No. 6218523  
GENERAL INFORMATION:  
APPLICANT: French, Cynthia K.  
APPLICANT: Schneider, Patrick A.  
APPLICANT: Yamamoto, Karen K.  
TITLE OF INVENTION: Prostate Cancer-Specific Marker  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/036,315  
FILING DATE: 06-MAR-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/047,811  
FILING DATE: 15-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/041,246  
FILING DATE: 07-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Storella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 018002-000210US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..120  
OTHER INFORMATION: /note="synaptotagmin "A" internal



Best Local Similarly 43.38; Pred. NO. 1.88

; Patent No. 5948664

Fri Dec 14 10:32:11 2001

us-09-680-121-2.rat

Page 8

```

; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: NO. 5948664el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-609-049A-16

```

```

Query Match      8.3%; Score 182.5; DB 2; Length 140;
Best Local Similarity 32.8%; Pred. No. 5e-11;
Matches 39; Conservative 29; Mismatches 48; Indels 3; Gaps 2;

QY 154 LGTLFESLVEYNEERKAFVNIKEARGLPAMDQSMSPYIKMTLP--EKHKVYKTRVL 211
    ||| ||| : : : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB 11 LGALEFSLHYDQNSSIHCTIIKAKGLKPMDSNGL-ADPYVKHLPLPGASKSNKLTRTL 69
    | : : : || : : : || : : : || : : : || : : : || : : : || : : : ||
QY 212 RKTLDPADEFETFYGIPTQIOELALHFTILSFDRSRDIIIGEVLIPLSGIELSECK 270
    | : : : || : : : || : : : || : : : || : : : || : : : || : : : ||
DB 70 RNTNRPIWNETLVYHGITDEDMQKRTLRLISVCDEDKFGHNEFIFGTRFSLKIKLPNQK 128
    | : : : || : : : || : : : || : : : || : : : || : : : || : : : ||

```

```

RESULT 15
US-09-170-996-16
; Sequence 16, Application US/09170996
; Patent No. 6291220
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: No. 6291220el PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/170,996
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,049
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 2307K-063700US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-170-996-16

```

```

Query Match      8.3%; Score 182.5; DB 4; Length 140;
Best Local Similarity 32.8%; Pred. No. 5e-11;
Matches 39; Conservative 29; Mismatches 48; Indels 3; Gaps 2;

QY 154 LGTLFESLVEYNEERKAFVNIKEARGLPAMDQSMSPYIKMTLP--EKHKVYKTRVL 211
    ||| ||| : : : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB 11 LGALEFSLHYDQNSSIHCTIIKAKGLKPMDSNGL-ADPYVKHLPLPGASKSNKLTRTL 69
    | : : : || : : : || : : : || : : : || : : : || : : : || : : : ||
QY 212 RKTLDPADEFETFYGIPTQIOELALHFTILSFDRSRDIIIGEVLIPLSGIELSECK 270
    | : : : || : : : || : : : || : : : || : : : || : : : || : : : ||
DB 70 RNTNRPIWNETLVYHGITDEDMQKRTLRLISVCDEDKFGHNEFIFGTRFSLKIKLPNQK 128
    | : : : || : : : || : : : || : : : || : : : || : : : || : : : ||

```

Search completed: December 12, 2001, 10:54:43  
Job time: 54 sec



1.  
2.  
3.  
4.



OM of: US-09-680-121-2 to: GenEmbl:\* out\_format : pfs

Date: Dec 12, 2001 11:33 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

#### Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/gen2.1/USPto.spool/US09680121/runat.12122001.105512.22734/app\_query.fasta.1.488  
-DB=genEmbl -QWMT=fastap -SUFFIX=p2n.rge -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000  
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NOR=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US09680121 -CGN1\_1\_0 -NCP0=6 -ICPU=3 -LONGLOG -NO\_XLPEXY  
-WAIT -THREADS=1

#### Search information block:

Query: US-09-680-121-2  
Query length: 425  
Database: GenEmbl:\*  
Database sequences: 1472140  
Database length: -341344837  
Search time (sec): 1640.590000

#### score\_list:

Sequence	Strd Orig	ZScore	Escore	len	Documentation
gb.pr:AF299075	+ 2201.00	3224.36	2.5e-171	1671	AF299075 Homo sapiens synaptr
gb.pac:ARI46667	+ 2201.00	3215.74	7.7e-171	1671	ARI46667 Sequence 1 from paten
gb.pr:AB037763	+ 2201.00	3215.69	7.7e-171	1671	AB037763 Homo sapiens mRNA for
gb.ro:RARSFI	+ 2006.00	2935.44	2.2e-155	1576	L38247 Rattus norvegicus (clon
gb.ro:RNU14398	+ 2006.00	2935.44	2.2e-155	1576	U14398 Rattus norvegicus syna
gb.ro:MMU10355	+ 1988.00	2802.23	3.1e-153	1576	U10355 Mus musculus BALB/c sy
gb.hlg:AC009266	+ 1624.00	2328.57	2.0e-121	1717	AC009266 Homo sapiens chrom
gb.hlg:AF001197	+ 1624.00	2327.51	2.3e-121	1905	AP001197 Homo sapiens chrom
gb.hlg:AC087507	+ 1624.00	2327.51	2.3e-121	1905	AC087507 Homo sapiens chrom
gb.hlg:AC091039	+ 1603.00	2298.46	9.5e-120	1593	AC091039 Homo sapiens chrom
gb.hlg:AC068125	+ 1595.00	2286.01	4.7e-119	1702	AC068125 Homo sapiens chrom
gb.pr:AB026808	+ 1167.50	1707.03	8.3e-87	1293	AB026808 Mus musculus mRNA for
gb.pr:AB004291	+ 1162.00	1690.54	6.9e-86	1293	BC004291 Homo sapiens. Similat
gb.ro:AF375445	+ 1160.50	1696.74	1.1e-86	1293	AF375445 Rattus norvegicus syn
gb.pr:AF027540	+ 1159.00	1688.60	8.8e-86	1235	AF027540 Homo sapiens CDNA FL
gb.ro:AF000423	+ 1156.50	1684.44	1.5e-85	2426	AF000423 Rattus norvegicus sy
gb.hlg:AF181098	+ 757.00	1094.52	1.1e-52	3067	AF181098 Drosophila melanogast
gb.hlg:AC017221	+ 689.50	961.54	2.8e-45	83824	AC017221 Drosophila melanogast
gb.in:AC008029	+ 689.50	953.92	7.4e-45	176991	AC008029 Drosophila melanogast
gb.in:AB003672	+ 689.50	948.49	1.5e-44	301601	AB003672 Drosophila melanogast
gb.hlg:AL356483	+ 677.50	938.47	5.6e-44	146316	AL356483 Homo sapiens chrom
gb.hlg:AL139128	+ 676.50	935.58	7.8e-44	163983	AL139128 Homo sapiens chrom
gb.in:AB044144	+ 645.50	932.45	1.2e-43	2544	AB044144 Halocynthia roretzi m
gb.ro:AB026804	+ 637.50	913.53	1.3e-42	1212	AB026804 Mus musculus mRNA for
gb.ro:AF336854	+ 637.50	913.53	1.3e-42	1212	AF336854 Rattus norvegicus syn
gb.ro:RNU20106	+ 627.00	908.21	2.6e-42	1900	U20106 Rattus norvegicus syna
gb.in:D63797	+ 623.50	901.66	6.0e-42	2180	D63797 Lolio pealei mRNA for
gb.in:LP5VNTGM	+ 621.00	903.44	4.8e-42	1277	X72386 L. pealei mRNA for syna
gb.ro:AF336856	+ 615.50	893.29	1.8e-41	1572	AF336856 Rattus norvegicus syn
gb.ro:AF336857	+ 615.50	893.23	1.8e-41	1572	AF336857 Rattus norvegicus syn
gb.ro:AF336858	+ 615.50	892.41	2.0e-41	1704	AF336858 Rattus norvegicus syn
gb.pr:AF038515	+ 615.50	891.75	2.1e-41	1818	AF038515 Homo sapiens synaptr
gb.ro:AF336859	+ 615.50	891.65	2.2e-41	1836	AF336859 Rattus norvegicus syn
gb.ro:AF336860	+ 615.50	891.13	2.3e-41	1932	AF336860 Rattus norvegicus syn
gb.ro:AF336853	+ 615.50	890.49	2.4e-41	1959	AF336853 Rattus norvegicus syn
gb.ro:AF336855	+ 615.50	890.46	2.5e-41	2064	AF336855 Rattus norvegicus syn
gb.in:AF331039	+ 613.00	889.87	2.7e-41	2187	AF331039 Manduca sexta synaptr
gb.in:ACU03125	+ 606.50	881.05	8.5e-41	1417	U03125 Aplysia californica syn

gb.ro:MUSSYNB + 605.00 875.99 1.6e-40 1876 D3793 Mouse mRNA for synap  
gb.ro:AF257304 + 605.00 871.92 2.7e-40 2795 AF257304 Mus musculus synap  
emb.pat:E11219 + 604.00 874.52 2.0e-40 1876 E11219 Mouse cDNA encoding  
gb.ro:RATSYNII + 604.00 870.88 3.1e-40 2681 M64488 Rat synaptrcagmin II  
gb.ro:RNU20105 + 603.50 873.92 2.1e-40 1850 U20105 Rattus norvegicus sy

seq\_name: gb.pr:AF299075

seq\_documentation block: 1671 bp mRNA PRI 04-DEC-2000

LOCUS AF299075 Homo sapiens synaptrcagmin IV mRNA, complete cds.

DEFINITION AF299075.1 GI:11528336

VERSION AF299075.1

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1671)

Ferguson,G.D., Chen,X.-N., Korenberg,J.R. and Herschman,H.R.

The Human Synaptrcagmin IV Gene Defines an Evolutionary Break Point

between Syntelic Mouse and Human Chromosome Regions but Retains

Ligand Inducibility and Tissue Specificity

J. Biol. Chem. 275 (47), 36920-36926 (2000)

10938284

2 (bases 1 to 1671)

REFERENCE Ferguson,G.D.

Direct Submission

Submitted (24-AUG-2000) Pharmacology, University of Washington,

Health Sciences J681, Seattle, WA 98195, USA

Location/Qualifiers

1. 1671

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="18"

/map="18q12.3"

94. 1371

/codon\_start=1

/product="synaptrcagmin IV"

/protein\_id="AK637229.1"

/db\_xref="GI:11528337"

/translation="MAPITTSREDEPDIPTVVGIFSAFLVFTVSLFAMICORRKS

SKTPPYKRVHVLKGVDIAPENILSKKKGADKNKKPAPKNSLHLEKRLN

GNPKPTKLVGSPSDLENATPKFLFEKEKESPELSKSTSLTSEKOEKGLTFES

LVNPERKAPVNIKARGLPMDQSDMPYIKRTTLPKKKKHYKTKRLLDPA

PBETFTYPTPTIOELALHTTILTSDFPRSRVDITGEVILPLSGISEGKLMRE

IKRINRKSSGREGELISLCYSTNTFTLVVAKARLKRSDVSGISDLYVKNLTHA

KKRISKKTGHVKCTPNAAVFNELFVFDICFEGLEDISVEFLVLDSESGSRNEVIGOLV

LGAACGTGCEHMKETCDYPRQIAKWHVLCDSG"

BASE COUNT 513 a 330 C 372 G 456 T

ORIGIN

alignment\_scores:

Quality: 2201.00 Length: 425

Ratio: 5.179 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-680-121-2 x AF299075 ..

Align seg 1/1 to: AF299075 from: 1 to: 1671

1 MetAlaPoiIeThrTnrSerArgLugLubPheAspGluIleProThrVa 17

|||||

94 ATGGCTCGATGACACACCGCGAAGAAATTTGATGAAATCCACAGT 143

|||||

17 lValGlyLePheSerIaPheGlyLeuValPheThrValSerIeuphea 34

|||||

144 GGTGGGATCTTCAGTGCATTTGGCTTGCTTCAAGTCTCTCTTTG 193

|||||

34 lATPILecyScGlnArgLysSerSerIySerAsnLysThrProPro 50

|||||

194 CATGATCTGCTGTCAGAGAAATCATCCAAAGTCTAACAAGACTCCTCCA 243  
 51 TYRFLYSRHEVALHISVALLEULYSGLYVALASPRIETYPROGLYASPLE 67  
 244 TACAGATTGTGCANTGCTTAAAGGAGTGCATATTACCTGAAAACCT 293  
 67 uasnserylsylsypheglialaaspasplysasngluvalylsasnl 84  
 294 AAATGCAAAAAGAGTTTGAGCAGATGATTAATAAGTAAGTAAGTAATA 343  
 84 yspcoalaValprolysasnserserleuHISleuaspleugluysarasp 100  
 344 ACCCAGCTGTGCCAAAGAAATTCATTCATCTGGATCTGAAAAGAGAGAT 393  
 101 LeuaasnglyasnphetrolysthrAsnleuLysProglySerProSerAs 117  
 394 CTCATATGGCAATTTTCCCAAAACCMACCTCAACCTGCAGCTCTTCTGA 443  
 117 pleuGLuasnaIathrProlyslsleupheleugluGLyGLyGLyGLySer 134  
 444 TCTGGAGATGCACACCCGAAAGCTCTTTTGAAGGGGAAAAAGAGTCAG 493  
 134 aLserProgluSerleuLysSerSerSerSerSerSerSerSerSerSer 150  
 494 TTTCCCTGAGAGTTTAAAGTCAGCAGCTTCCTTACTTCAAGAGAGAAA 543  
 151 GLNGLUYLSLEUGLYTHRleuphepserleuGLUTYRAsnPhesGLUAR 167  
 544 CAAAGAAAGCTGGGAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 593  
 167 GLYSLAaheValValasnIlelysgluAaargGLYleuProalaMetA 184  
 594 AAACCATTTTGTGCTCATATCAAGAGAACCCGGCTTGGCCAGCCATGG 643  
 184 spGLUGLISerMetThrSerAspProtyrIleLysMetThrIleleuPro 200  
 644 ATGACACATGCATGACCTCTGACCATATATCAAAATGACGATCTCCCA 693  
 201 GLUYSLYSHISLYSVALYsthrArGValleuArGLYsthrleuAspPr 217  
 694 GAGAAAGACATAAAGTGAAGAACTAGAGTCTGAGAAAACCTTGATCC 743  
 217 oAlaIaheaspGLUthRphetRphetRphetRphetRphetRphetRph 234  
 744 ACCTTTGATGAGACCTTACATTCATTCATTCATTCATTCATTCATTC 793  
 234 InGLuleuAlaLeuHISPhetThrIleleuSerPheaspRpheserArg 250  
 794 AAGATTTGGCTTGCCTCACTCACAATTTTGAAGTTTGACAGGTTTCAAG 843  
 251 AspaPPIleIleGLYGLYValleuIleProleuSerGLYIleGLuleuSe 267  
 844 GATGATATCATTTGGGAGATTTCTAATTCCTCTCTGGGAAATTAATATC 893  
 267 rGLUGLYSmetleuMetAsnaArgLUleIleLysarGAsnValArgL 284  
 894 TGAAGGAAAAATGTTAATGAATGAAGAGATCAACAGAAAATGTTTGAAG 943  
 284 ySserSerGLYArgGLYGLuleuLeuIleSerleuLysArgIleSerThr 300  
 944 AGTCTTCAGGAGCGGGAGTGTACTGATCTCTCTGCTATCAGTCCACC 993  
 301 ThrAsnThrleuThrValValleuLysAlaArgHISleuProLysSe 317  
 994 ACAAACTACTTAAGTGTGTCTTAAAGCTGCAGCATCTGCCCTAATC 1043  
 317 rAspValSerGLYleuSerAspProtyrValLysValasnleuThrHSA 334  
 1044 TGAATGTGCGGACTTTCAGATCCCATGTCAAAAGTGAACCGTACCATG 1093  
 334 lALysLYSArgIleSerLysLysLysThrHISValLysLysCysThrPro 350  
 1094 CCAAAAGAGATCTCCAGAGAGAGATCATGTGTAAGAAATGACCCCC 1143

351 AsnAlaValaPheasnGLuleupheValaPheaspRIleProCysGLYlyLe 367  
 1144 AATGCACTGTTCATATGAGCTGTTTGTCTTGTATTCCTTGTGTAGGGCT 1193  
 367 uGLUAsPIleSerValaGLUphleuValleuAspSerGLUArgGLYSerA 384  
 1194 TGAAGATATAGGTGTGAATTTTGTGTTGATTCGAAGGGGGTCCC 1243  
 384 rGAsnGLUValIleGLYGLuleuValleuGLYAlaAlaIaGLYlyThr 400  
 1244 GAAATGAGTAAATCGGGCAGTATGCTTGGGTGCAGCAGCAAGAGACT 1293  
 401 GLYGLYGLUthIstRplysGLUleCysAspTYRProArGArgIleAla 417  
 1294 GGTGAGAGCAGCTGAAAGAGATCTGTGACTACCCAGAGACAAATGTC 1343  
 417 aLYSTRPHISValleuCysAspGLY 425  
 1344 CAAGTGCACGCTCTCTGTGATGT 1368  
 seq\_name: gb\_pat:AR146667  
 seq\_documentation\_block: 3891 bp DNA PAT 08-AUG-2001  
 LOCUS AR146667  
 DEFINITION Sequence 1 from patent US 6218523.  
 ACCESSION AR146667  
 VERSION AR146667.1 GI:15109856  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 3891)  
 AUTHORS French,C.K., Schneider,P.A. and Yamamoto,K.K.  
 TITLE Prostate cancer-specific marker  
 JOURNAL Patent: US 6218523-A 1 17-APR-2001;  
 FEATURES  
 source 1..3891  
 BASE COUNT 1229 a 671 c 788 g 1203 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 2201.00 Length: 425  
 Ratio: 5.179 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000  
 alignment\_block:  
 US-09-680-121-2 x AR146667  
 Align seg 1/1 to: AR146667 from: 1 to: 3891  
 1 MetaIaProIleThrThrSerArgGLUleuPheaspGLUleProthryA 17  
 151 ATGGCTCCATCACCCACCGCGGAAAGAAATTTGATGAATCCCCACAGCT 200  
 17 lValaGLYIlePheSerAlaPheGLYleuValaPheThrValSerleuPheA 34  
 201 GGTGGGATCTTCAGTGCATTTGGCTGTCTTCACAGTCTCTCTCTTG 250  
 34 lATRIleCysCysGLNArgLysSerSerLysSerAsnLysThrProPro 50  
 251 CATGATCTGCTGTCAGAAATCATCAAGTGTACAGAGACTCCTCCA 300  
 51 TYRFLYSRHEVALHISVALLEULYSGLYVALASPRIETYPROGLYASPLE 67  
 301 TACAAGTTTGTGCATGCTCTTAAGGAGATTGATTTTACCTGAAAACCT 350  
 67 uasnserylsylsypheglialaaspasplysasngluvalylsasnl 84  
 351 AAATGCAAAAAGAGTTTGAGCAGATGATTAATAAGTAAGTAAGTAATA 400

```

84 yspProAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
   |||||||
401 AGCAGACTGTGCCAAAGAAATTCATCTGATCTGATCTTGTAAAAAGAGAT 450
   |||||||
101 LeuAsnGlyAsnProLysThrAsnLeuLysProGlySerProSerAs 117
   |||||||
451 CTCATATGGCAATTTTCCCAAAACCAACCTCAAACTGGCAGTCTTCTGTA 500
   |||||||
117 pLeuGluAsnAlaThrProLysLeuPheLeuGluGlyGluLysGlySer 134
   |||||||
501 TCTGGAGAAATGCACACCCCGAAGCTCTTTTGAAGGGGAAAAAGATCAG 550
   |||||||
134 alSerProGluSerLeuLysSerSerThrSerLeuThrSerGluGluLys 150
   |||||||
551 TTTTCCCTGAGAGTTTAAAGTCCAGACATTCCTTACTTCCAGAGAGAAA 600
   |||||||
151 GluGluLysLeuGluLysThrLeuPhePheSerLeuGluLysThrAsn 167
   |||||||
601 CAAGAGAGCTGGGAGACTCTCTTCTCTCTTGAATACAACTTCAGAGAG 650
   |||||||
167 glysAlaPheValValAsnLleLysGluAlaArgGlyLeuProAlaMet 184
   |||||||
651 AAAAGCATTTGTGTCATATATCAAGAGACCCCGTGGCTGGCCAGCATGG 700
   |||||||
184 spGluGlnSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
   |||||||
701 ATGAGCAGTCGATGACCTCTGACCATATATCAAAATGACGATCTCCCA 750
   |||||||
201 GluLysLysHisLysValLysThrArgValLeuArgLysThrLeuAsp 217
   |||||||
751 GAGAGAGAGCATGAAGTGAAGAACTAGAGTCTGAGAAAAACCTTGATCC 800
   |||||||
217 alaIaPheAspGluThrPheThrPheTyrGlyIleProTyrThrGlnIle 234
   |||||||
801 AGCTTTTGATGAGACCTTTATCATTTATGGATACCTTACACCAATTC 850
   |||||||
234 InclLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
   |||||||
851 AAAAAATGGCTTGCACCTTCAATTTTGAGTTTGACAGTTTTCAGA 900
   |||||||
251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeu 267
   |||||||
901 GATGATATCATTTGGGGAAGTTCTAATTCCTCTCGGGAATGAATATTC 950
   |||||||
267 rglGluLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284
   |||||||
951 TGAAGGAAAAATGTTATGATAGAGATCATCAAGAAATGTTATGGA 1000
   |||||||
284 ysserSerGlyArgGlyGluLeuLeuIleSerLeuGlyTyrGlnSerThr 300
   |||||||
1001 AGTCTTCAGAGACGGGTGAGTTACTGATCTCTCTCTCTATCAGTCCACC 1050
   |||||||
301 ThrAsnThrLeuThrValValValLeuLysAlaArgHisLeuProLys 317
   |||||||
1051 ACAAAACCTCTAATCTGCTGTCTTAAAGCTCGACATCTGCCTAATTC 1100
   |||||||
317 rAspValIleSerGlyLeuSerAspProTyrValLysValAsnLeuThrHis 334
   |||||||
1101 TGAATGTCCTCGACCTTCAGATCCCTATGTCAAGTGAACCTGTACCAT 1150
   |||||||
334 lAlaLysArgGlyIleSerLysLysLysThrHisValLysLysCysThrPro 350
   |||||||
1151 CCAAAAAGAAATCTCCAGAGAGACATCATGTGAAGAAATGACCCCC 1200
   |||||||
351 AsnAlaValPheAsnGluLeuPheValPheAspIleProCysGluGlyLe 367
   |||||||
1201 AATGCAAGTGTTCATATGAGCTGTGTCTTGAATATCTGTGAGGGCCT 1250
   |||||||
367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgGlySer 384
   |||||||
1251 TGAAGATATATGATGATTTTGGTTTGGATCTGAAAAGGGGGTCC 1300
   |||||||
384 rgaSngluValIleGlyGlnLeuValLeuGlyAlaAlaIaGluGlyThr 400

```

```

|||||
1301 GAAATGGGTAATCGGGCACTGATAGTCTGGGTGCACAGAGAAAGACT 1350
   |||||||
401 GlyGlyGluHisThrLysGluIleCysAspTyrProArgArgLleAl 417
   |||||||
1351 GTTGAGAGACGACTGAAGAGATCTGTACTACCCAGAGACAAATTC 1400
   |||||||
417 alYsTrpHisValLeuCysAspGly 425
   |||||||
1401 CAAGTGGCACGTGCTGTGATGT 1425
   |||||||
seq_name: gb_pr:AB037763

seq_documentation_block:
LOCUS AB037763 3910 bp mRNA PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1342 protein, partial cds.
ACCESSION AB037763
VERSION AB037763.1 GI:7243064
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_lib:pbLuescriptII SK plus
clone:fj00418.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE Nagase,T., Kikuno,R., Ishikawa,K.I., Hitosawa,M. and Ohara,O.
1 (sites)
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
20181126
REFERENCE 2 (bases 1 to 3910)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) to the DDBJ/EMBL/GenBank databases. Osamu
Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology;
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:cdna@fokkazusa.or.jp, URL:http://www.kazusa.or.jp/huge/,
Tel:+81-438-52-3913, Fax:+81-438-52-3914)

FEATURES
source
1..3910
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fj00418"
/clone_lib="pbLuescriptII SK plus"
/tissue_type="brain"
184..1464
/gene="KIAA1342"
<184..1464
/gene="KIAA1342"
/protein_id="BAA92580.1"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1342 protein"
/protein_id="BAA92580.1"
/db_xref="GI:7243065"
/translation="KMADITTSREEDFELPTVVGIFSAFLGVFTVSLPAMICORRKS
KSNKTPYKFEVLKGVDIYENLNSKKKFGADKNEVKRAVPKNSJHDLERDL
NGNPKTNLKCPSPDIDENAPKLEEGKEVSPESLKSSTLSSEKOEKLTGLTF
SLEYNPERKAPVYNKEARGDLPAMDOSHTSPYIKMTILPEKKHKVRYRKTLDP
AFDEFTFYGIPTQIQLALHFTLLSDFERSDIDIGVLLPLSGIEISEKMLNR
ELIKRNVKSSSGRGLLSICVOSTTNTLTJVVVLARHLPKSDVGLSDPYKVNLYH
AKRIRSKKTHTVKKCTPNVAFNELFVFDLPCGLDIDISVEFLVLDSEKRSRNEVIGOL
VGAAGAEKGSGHMKEDYPRROIAKMHVLCDG"

BASE COUNT 1221 a 679 c 799 g 1211 t
ORIGIN

alignment_scores:
Quality: 2201.00 Length: 425
Ratio: 5.179 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

```

US-09-680-121-2 x AB037763 ..  
Align seg 1/1 to: AB037763 from: 1 to: 3910

```

1 MetAlaProIleThrThrSerArgGluGluPheAspGluIleProThrVal 17
187 ATGGCTCCGATCCACCACCGAGAAATTTGATGAATATCCACAGT 236
17 ValAllyIlePheSerAlaPheCtyLeuValPheThrValSerLeuPhe 34
237 GGTGGGATCTTCAGTCATTTGGCTGCTTCACAGCTCTCTCTTTG 286
34 IATPILecCysCysGlnArgLysSerSerLysSerAsnLysThrProPro 50
287 CATGGATCTGCTGCAGAGAAATTCATCCAGCTTACCAAGACTCTCTCA 336
51 TyrLysPheValHisValLeuLysGlyValAspIleTyrProGluAsnLe 67
337 TACAGATTGTGCATGTGCTTAAGCAGATTGATATTTACCTGAAACCT 386
67 uAsnSerLysLysPheGlyAlaAspAspLysAsnGluValLysAsnL 84
387 AAATGCAAAAAGAGTTTGGACGATGATAAATGAAGTAAAGATA 436
84 YSPROAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
437 AGCCAGCTGTCCCAAGAAATTCATTCATCTGATCTTGAAAAGAGAT 486
101 LeuAsnGlyAsnPheProLysThrAsnLeuLysProGlySerProSerAs 117
487 CTCATGGCAATTTCCCAACCAACCTAACCTGGCAGCTCTCTGGA 536
117 PLeuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGluServ 134
537 TCTGAGAGATGCAACCCCGAAGCTCTTTTAAAGGGAGAAAAGGTCAG 586
134 aLSerProGluSerLeuLysSerSerThrSerLeuThrSerGluGluLys 150
587 TTTCCCTCGAGAGTTTAAAGTCCAGCATCTCCCTTACCTTCAGAGAGAA 636
151 GlnGluLysLeuGlyThrLeuPhePheSerLeuGluLysArgAsnPheGlu 167
637 CAAGAGAGCTGGGAACTCTCTCTCTCTTGAATGATCAACTTCGAGAG 686
167 GlySerAlaPheValValAsnLysGluValArgGlyLeuProAlaMet 184
687 AAAAGCATTTGTGTCAATTCAGAGAACCCGCTGCTTGCACCATG 736
184 SPGLuGlnSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
737 ATGAGCAGCTGCATGACCTCGACCCATATATATAATACATCATCTCCA 786
201 GlnLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217
787 GAGAGAGACATAAGAGAACTAGAGCTGAGAGAAAACCTTTGGATCC 836
217 AlaPheAspCyluThrPheThrPheTyrGlyLeuProTyrThrGlnIle 234
837 AGCTTTGATGAGAACCTTACATCTATGAGTACCTACCAACCCAAATCC 886
234 InGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
887 AAGAAATGGCCCTTGCACTTCACAAATTTGAGATTGACAGCTTTCAAGA 936
251 AspAspIleIleGlyGluValLeuLeuProLeuSerGlyIleGluLeuSe 267
937 GATGATATCATTTGGGAGAGTTCTAATCTCTCTCGGGAATTTGATATTC 986
267 rGluGlyLysMetLeuMetAsnArgGluIleLeuLysArgAsnValArgL 284
987 TGAAGGAAAAATGTTAATGATAGAGATCATCAAGAGAAATTTTNGGA 1036
284 ySerSerGlyArgGlyGluLeuLeuIleSerLeuCysTyrGlnSerThr 300

```

```

1037 AGCTTCAGGACGGGTGAGTTACTGATCTCTCTCTCTATCATGCTCAC 1086
301 ThrAsnThrLeuThrValValValLeuLysAlaArgHisLeuProLysSe 317
1087 ACAACAGCTTAACTAGTGTGTCTTAAAGCTCGACATCTGCTTAAATC 1136
317 rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuTyrHisA 334
1137 TGATGTGTCGGGACTTTCAGATCCCTATGTCMAAGTGAACCTTACATG 1186
334 lAlLysLysArgIleSerLysLysThrHisValLysLysCysThrPro 350
1187 CCAAAAAGAGAAATCTCCAAAGAGAGACTCATGTGAGAAATGCACCC 1236
351 AsnAlaValPheAsnGluLeuPheValPheAspIleProCysGluGlyLe 367
1237 AATGCACTGTTCATGAGCTGTGTGCTTTCATATCTCTGTCAGGGCT 1286
367 uGluAspIleSerValGluPheLeuValLeuAspSerCyluArgLysArg 384
1287 TGAAGATATAGCTTGAATTTTGGTTTGGATTCGAAAGGGGCTCC 1336
384 rGAsnGluValIleGlyGlnLeuValLeuGlyAlaAlaIleGluGlyThr 400
1337 GAATAGAGGTAATCGGCGAGTTAGCTTGGGTGCGACGACGAGAGAACT 1386
401 GlyGlyGluHisTrrPlyGluIleCysAspTyrProArgArgGlnIleAl 417
1387 GTTGAGAGACACTGGAAGAGATCTGTGACTACCCAGACCAAAATTC 1436
417 aLysTrrPHisValLeuCysAspGly 425
1437 CAAGTGCAGCTGCTCTGTGATGAT 1461

```

seq\_name: gb\_ro:RATRSI

seq\_documentation\_block:

LOCUS RATRSI 1576 bp mRNA ROD 24-MAY-1995  
DEFINITION Rattus norvegicus (clone PCR2) synaptotagmin IV mRNA, complete cds.  
ACCESSION L38247.1 GI:598376  
VERSION 1.0  
KEYWORDS immediate early gene; synaptotagmin; synaptotagmin IV;  
SOURCE Rattus norvegicus (strain New England Deaconess Hospital) (clone  
library: lambda ZAP-PCR2+ library (L. Vician and I.K. Lim))  
PHEOCHROMOCYTOMA cDNA to mRNA.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1576)

REFERENCE  
AUTHORS Herschman, H.R.  
TITLE Synaptotagmin IV is an immediate early gene induced by  
depoloarization in PC12 cells and in brain  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 92 (6), 2164-2168 (1995)  
MEDLINE 95199312  
FEATURES  
source  
1. location/Qualifiers  
1.1576  
/organism="Rattus norvegicus"  
/strain="New England Deaconess Hospital"  
/db\_xref="taxon:10116"  
/cell\_line="PC12"  
/clone="PCR2"  
/clone\_lib="lambda ZAP PCR2+ library (L. Vician and I.K. Lim)"  
/tissue\_type="pheochromocytoma"  
1.239  
240..1517  
/codon\_start=1  
/product="synaptotagmin IV"  
/protein\_id="AA067327.1"  
/db\_xref="GI:598377"

5'UTR  
CDS



/translation="MAPITTSHEDEIPITVVGFSAFLVFTVSLFAMICCRRSK  
 SNKTPYKRVHLYKGDIVPENLSSKKRGDSKAKKALPNLSLIDLKRLN  
 GNPRTNPKAGSSSDLENTYPLPETEREANSPELSKSTLSTSEKEXICLPLS  
 LEYNEKFAVYNIKEAOGLPAMDOSMTSDPYIKTTLPEKHKVKTIVLKRITLDPV  
 FDETFYGVVPIQELSLHFTVLSDFRSBDVIGELVPLSGLEISDGMKLMRE  
 IIRNAKSSGSELVSLCYOSTNTLLVVKAKHPLKSDVSGISLSDYVKNLYHA  
 KRISKKTTHVKCTPNNAVENLEFVEDICELESIEVFLVLDSESGSRNEVIGRLV  
 LGATGSGSGGGMKEICDFPRROIAKMHMLCDG"  
 1514 1576  
 3'UTR  
 BASE COUNT 420 a 394 c 384 g 378 t  
 ORIGIN

alignment\_scores:  
 Quality: 2006.00 Length: 425  
 Ratio: 4.834 Gaps: 0  
 Percent Similarity: 97.647 Percent Identity: 89.882

alignment\_block:  
 US-09-680-121-2 x RATRSI ..

Align seg 1/1 to: RATRSI from: 1 to: 1576

1 MetalabroliethrhrSerArgLugluPheaspGluileProthVa 17  
 |||||  
 240 ATGGCTCTATACACACAGCCGGTGGAATTCGATGAATTCACAGT 289  
 17 lValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPhe 34  
 |||||  
 290 GGTGGGATCTTCAGTCTTTGGCTGGCTTCACTGTCCTCTCTTG 339  
 34 lATPPIleCySGlnArgLysSerSerLysSerAsnLysThrProPro 50  
 |||||  
 340 CCGTGGATCTGTCGACAGAGATCCGTAAGTCCACAGACCTCTTCA 389  
 51 TyrIysPheValHisValLeuLysGlyValAspIleYrProGluAsnLe 67  
 |||||  
 390 TACAGATTGTGCACGTCCTTAAAGAGATTATCTACCCAGAAACCT 439  
 67 uAsnSerLysLysLysPheGlyAlaAspAspLysAsnGluValLysAsnL 84  
 |||||  
 440 AAGTAGCAAAAGAGATTGGAGAGATGACAGAGTGAAGCAAAAGCGTA 489  
 84 ysProAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgasp 100  
 |||||  
 490 AAGCAGCTTCCCTAACCTTCCCTGCACCTTGATCTGAGAGAGGAAAC 539  
 101 LeuAsnGlyAsnAlaPheProLysThrAsnLeuLysProGlySerProSerAs 117  
 |||||  
 540 CTCATGTGCATTTCCCAAAACCAACCCCAAGAGCTGCAGCTTCTGA 589  
 117 pLeuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGluSerV 134  
 |||||  
 590 TCTGGAATATGTCAACCCCAAGCTTTCGAGAGCGGAAAGAGGCGC 639  
 134 alSerProGluSerLeuLysSerSerThrSerLeuThrSerGluGluLys 150  
 |||||  
 640 TCTCCCTCGAGAGCTTGAGATCCAGACCTTCCCTCATTCCAGAGAGAAA 689  
 151 GlnGluLysLeuGlyThrLeuPhePheSerLeuGluLysArgAsnPheGluAr 167  
 |||||  
 690 CAAGAGAGCTGGGACACTCTTCTGTCTAGAGTACCACTTGAGAAA 739  
 167 gLysAlaPheValAlaAsnIleLysGluAlaArgLysLeuProAlaMetA 184  
 |||||  
 740 GAAAGCATTTGTGTGAACATCAAGAGAGCCAGGCTTACAGCATGAG 789  
 184 spGluGluSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200  
 |||||  
 790 ATGACAAATTCATGACCTTGACCATCAATCAATGATCAATCTTACCA 839  
 201 GluLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217  
 |||||

840 GAGAGAAGACACAAAGTGAAGAAACAGAGAGTGTGAGAGAGAGCTGAGACC 889  
 217 oAlaPheaspGluThrPheThrPheTyrGlyIleProTyrThrGlnIleG 234  
 |||||  
 890 CGGTGTGGACAAACCTTCACATTCGTAGGGGCTCCCTTATCCACACACC 939  
 234 lngluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250  
 |||||  
 940 AAGAGCTCTCCCTCAGCTTCCAGCTCCGAGATTTGACAGAGTTCTCAAGA 989  
 251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeuSe 267  
 |||||  
 1040 AGACGGGAAATGTTAATGACACAGAGATCATCAAGAGAAATCTTAAGA 1089  
 267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284  
 |||||  
 284 ysSerSerGlyArgGlyLysLeuLeuIleSerLeuCyTyrGlnSerThr 300  
 |||||  
 1090 AGTCTTCGACAGGGGTGACTTCTGGTCTCCCTCTGTATCAGTCCACT 1139  
 301 ThrAsnThrLeuThrValValValLeuLysAlaArgHisLeuProLysSe 317  
 |||||  
 1140 ACAAACAGCTCAGCTGCTGTTGTTAAAGCGGACCTACCGAAATC 1189  
 317 rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuYrHisA 334  
 |||||  
 1190 TGATGTCTGACCTTCAGATCCCTACGTCAAAGTGAACCTGTACCAATG 1239  
 334 lAlsLysArgIleSerLysLysThrHisValLysLysCysThrPro 350  
 |||||  
 1240 CCAAGAAAGATCTCTAATAAGAAAGACTACGCTAATAAGGACCTCC 1289  
 351 AsnAlaValPheAsnGluLeuPheValPheAspIleProCysGluGlyLe 367  
 |||||  
 1290 AACGAGAGTTCACAGAACTGTTGTCTTATATCTCTTGAGAGACT 1339  
 367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgLysSerA 384  
 |||||  
 1340 TGAGAAATTAAGTGTGAATTTTATGTTTGGATTCATAAAGGAGATCCC 1389  
 384 rGAsnGluValIleGlyLysLeuValLeuGlyAlaAlaIleGluLysThr 400  
 |||||  
 1390 GAATAGAGGTGATCGGGGCTGGTTCCTGGGTCACAGACAGAAAGAGAT 1439  
 401 GlyGlyGluHisThrPylsGluIleCysAspTyrProArgGlnIleAl 417  
 |||||  
 1440 GGTGGGGGACACTGGAAAGAGATCTGTGACTTCCCCAGAGAGCAAAATTGC 1489  
 417 aLysThrHisValLeuCysAspGly 425  
 |||||  
 1490 TPAAGTCATATGCTCTGTGATGTT 1514  
 seq\_name: gb\_ro:RNUI4398  
 seq\_documentation\_block:  
 LOCUS RNUI4398 2060 bp mRNA 21-JUN-1995  
 DEFINITION Rattus norvegicus synaptotagmin IV homolog mRNA, complete cds.  
 ACCESSION U14398  
 VERSION U14398.1 GI:550453  
 KEYWORDS  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 2060)  
 AUTHORS Ullrich,B., Li,C., Zhang,J.Z., McMahon,H., Anderson,R.G.,  
 Geppert,M. and Sudhof,T.C.  
 TITLE Functional properties of multiple synaptotagmins in brain  
 JOURNAL Neuron 13 (6), 1281-1291 (1994)  
 MEDLINE 95085772

```
REFERENCE 2 (bases 1 to 2060)
AUTHORS Sudhof,T.C.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1994) Thomas C. Sudhof, Howard Hughes Medical
Institute, University of Texas Southwestern Medical Center, 5323
Harry Hines Boulevard, Dallas, TX 75235-9050, USA
FEATURES
SOURCE location/Qualifiers
1..2060
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone_lib="Strategene lambda ZAP adult rat brain cDNA
library"
/lisue_type="brain"
/dev_stage="adult"
268..1545
/note="similar to protein encoded by mouse synaptotagmin
IV, Genbank Accession Number U10355"
/product="synaptotagmin IV homolog"
/codon_start=1
/protein_id="AAA68519.1"
/db_xref="gi:550454"
/translaltion="MAPITTSRVEFDEIPYVGIFSAFLVFTVSLFAMICORSAS
SNKTPPYKRVHVLAKGVDIYPENLSKKFGDDKSAKKKALPNLSHLDEKRLDN
GNEPKNPAGSSDLENVTPKLEPETEREAVSPESLSTSEERKEKGLTFLPS
LEVNEPKAFVYVNIKEAGLPAMDQSMSTDPYIKMTIILPEKHKVTRVRLKTLDPY
FDDETFYGVGPYHIOELSHFTVLSFDRFSRDYDGEVLVPLSGIELSDPYKVNLYHA
IKRNKKSSSGGELVSLCYOSTNTLTVVILKARHLKPSDVSGLSDPYKVNLYHA
KRISKKTTHVKKCPNNAVEFELFDPIDPESIELEISVEFLVLDSESRNREVLGRLY
LQATGSGGGGHWKEIDFPFRROIKAHMLCDS"
BASE COUNT 559 a 474 c 480 g 547 t
ORIGIN
alignment_scores:
Quality: 2006.00 Length: 425
Ratio: 4.834 Gaps: 0
Percent Similarity: 97.047 Percent Identity: 89.862
Alignment_block:
US-09-680-121-2 x RNUI4398 ..
Align seg 1/1 to: RNUI4398 from: 1 to: 2060
1 MetAlaProIeIhThrSerArgGluGluPheAspGluIleProThrVa 17
|||||
268 ATGGCTCTATCACCACCGCGCTGGAAATCGATGAATCCACACAGT 317
|||||
17 lValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPheA 34
|||||
318 GGTGGGCATCTTCAGTCTTTGGCTGGCTTCACCTGTCCTCTCTTG 367
|||||
34 lAtPLeIcYcSgIaRgLySerSerLySerAsnLySthPrPro 50
|||||
368 COTGATGCTGCTGCCAGAGAATCCGTAAGTCCACAAGACTCTCTCA 417
|||||
51 TyrIlySPheValHisValLeuLySgLyValAspIleTyProGluAsnLe 67
|||||
418 TACAAGTTTGCGACGCTCTTAAGAGATTGATTCATCCCGAGAAACCT 467
|||||
67 uAsnSerLySgLySPheGlyAlaSPasPlySasngLlValIlySasnL 84
|||||
468 AAGTGCAGAAAGAGTTTGGAGAGATGACAGAGTGAAGCAAGCGCTA 517
|||||
84 ySPROAlaValProlYsAsnSerLeuHisLeuAspLeuGluLySArgAsp 100
|||||
518 AAGCGCTTGCTTAACCTTCCTCCACCTGATCTCGAGAAGGAGAAC 567
|||||
101 LeuAsnGlyAsnPhProLySthAsnLeuLySProGlySerProSerAs 117
|||||
568 CTCATGTGCATTTCCCAAAACCAACCCCAAGCTGCAGCTCTTCTGA 617
|||||
117 pLeuGluAsnAlaThrProLySLeuPheLeuGluGlyLySgIuserv 134
|||||
618 TCTGCAAAATCTCACCCCAAAAGCTCTTCCGGAGACGAAAGAGCCG 667
134 aLseRProGluSerLeuLySserSerThrSerLeuThrSerGluGlyLyS 150
668 TCTCCCTGAGAGCTTGAAGTCCAGACCTCCCTCACTCAAGAGAGAAA 717
151 GluGluLySLeuGlyThrLeuPheSerLeuGluTyTrAsnPhGluAr 167
718 CAAGAGAGCTGGGACCTCTCTTCTCTTAAGTACACACTTCAGGAA 767
167 gLySAlaPheValAlaSnIleLySgAlaArGlyLeuProAlaMeTa 184
768 GAAAGCATTTGTGTGAACATCAGAGAGCCAGGCGTTCACGACCATGG 817
184 sPGluGlnSerMetThrSerAspProTyTrIleLySmetThrIleLeuPro 200
818 ATGACACATCATGACCTCTGACCATCATCAAGATGACATCTTACCA 867
201 GluLySlyShIlySValLySthArgValLeuArGlySthLeuAspR 217.
868 GAGAGAGACCAAAAGTGAAGAACAGAGTCTGAGAGAACGCTGGACCC 917
217 oAlaPheAspGluThrPheThrPheTyGlyIleProTyTrThrGlnIleG 234
918 CGGTGTGACGAAACCTTCACATTCTATGGGGTCCCTTATCCCACTCC 967
234 InGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
968 AAGAGCTGCTCCGCACTTCACCGTCCGAGATTTCACAGSTTCTCAAGA 1017
251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeuSe 267
1018 GATGATGCTTGTGAGAGAGTCTCTCTCTTCACGAGATGAATATGTC 1067
267 rGluGlyLySmetLeuMetAsnArgGluIleIleLySArgAsnValArgL 284
1068 AGAGGGGAAATGTAAATGACACAGAGATCATCAGAGAAAGCTCAAGA 1117
284 ySserSerGlyArgGlyGluLeuLeuIleSerLeuCyStryGlnSerThr 300
1118 AGTCTCTGAGCGGGGTGAACCTTCGTCTCTCTCTGTATCAGTCCACT 1167
301 ThrAsnThrLeuThrValValLeuLySAlaArgHisLeuProLySse 317
1168 ACAAAACGCTCACTGTGTGTCTTAAGACCGCGCACCTACCGAAATC 1217
317 rAspValSerGlyLeuSerAspProTyTrValLySValAsnLeuTyTrHisA 334
1218 TGATGCTGTCGCACTTCAGATCCCTACGTCAAGTGAACCTGTACCATG 1267
334 lAlLySArGlyIleSerLySlySthHisValLySlyScySthPrPro 350
1268 CCAAGAGAGAAATCTCTAAAAAGAAAGACTCAGTGAAAGAGTCACTCCG 1317
351 AsnAlaValPheAsnGluLeuPheValPheAspIleProGlyGluLyLe 367
1318 AACGAGTGTCAAGAACTGTTGTCTTGTATTTCTCTGTAGAGTCT 1367
367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgLySerA 384
1368 TGAAGAAATGAAGGTTGAATTTTGTGTGATTCGAAAGGGGATCCC 1417
384 rGAsnGluValIleGlyIleLeuValLeuGlyAlaAlaIaGluGlyThr 400
1418 GAAATGAGTGTATCGGGGGGTGGTCTCGGTGTCACAGAGAGAAAGT 1467
401 GlYlGlyLuhIstRPlYgSgluIleCySAsPtyrProArGrgInIleAl 417
1468 GTGGGGGCACTGAGAGAGATCTGTGACTCCCGAGAGCAAAATTCG 1517
417 aLyStrPhIstValLeuCySAspGly 425
1518 TAAGTGCAATATGCTCTGTGATGGT 1542
```

```

seq_name: gb_ro:MMU10355
seq_documentation_block: 3992 bp mRNA ROD 23-AUG-1994
LOCUS MMU10355 Mus musculus BALB/c synaptotagmin 4 (syta) mRNA, complete cds.
DEFINITION U10355
ACCESSION U10355
VERSION U10355.1 GI:499125
KEYWORDS
SOURCE
  house mouse.
  Mus musculus.
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 3992)
  Hilbush, B.S. and Morgan, J.I.
  A third synaptotagmin gene, syt3, in the mouse
  Proc. Natl. Acad. Sci. U.S.A. 91, 8195-8199 (1994)
JOURNAL
  9436712
MEDLINE
  2 (sites)
  Perin, M.S., Fried, V.A., Mignery, G.A., Jahn, R. and Sudhof, T.C.
  Phospholipid binding by a synaptic vesicle protein homologous to
  the regulatory region of protein kinase C
  Nature 345 (6272), 260-263 (1990)
JOURNAL
  90238548
REFERENCE
  3 (bases 1 to 3992)
  Hilbush, B.S.
  Direct Submission
  Submitted (03-JUN-1994) Brian S. Hilbush, Roche Institute of
  Molecular Biology, 340 Kingsland St., Nutley, NJ 07110, USA
  Location/Qualifiers
FEATURES
  source
    1..3992
    /organism="Mus musculus"
    /strain="BALB/C"
    /db_xref="taxon:10090"
    /clone="p171"
    /tissue_type="brain"
    /gene="syt4"
    /gene="syt4"
    /gene="syt4"
    /gene="syt4"
    /product="synaptotagmin 4"
    /protein_id="AAA20971.1"
    /db_xref="GI:499126"
    /translation="MAPITTSRVEFDEIPTVYVIGISAFGLVTVSLFAMICCORBSAK
    SKRTPPKFVHYRLKVDITPELSSOKKRGDGKSEVKRTALPINLSLHLEKRDIN
    GNPFRANPAVAGSSSDLENVTPKLFETEEKANSPESLKSTLTSEKQEKGLTFLS
    LEYNEKAFVNIKEADGLPAMDEOSMTSDYIKWTILPEKHKRYKTVLKRKTLDPV
    FGETTFYGIPIPHIOELSLHPTVLSDFRSRDVYGEVLIPLSGIELSDGKMLMTRE
    IIKRAKKSSSGELLVSLCYOSTNTLTIVVVLKARHLPEKSDVSGISDPRYVKNLYHA
    KRRIKKRTHYKCTPNAYFNELFVDFICESELEISVFLVLDSESGSRNEYIGRLV
    LCAIYAGSGGGMKEICDEPRKQIAKMHMLCDG"
  misc-feature
    2157..2486
    /gene="syt4"
    /note="C2A repeat region, amino acids 148-257, internal
    repeat homologous to the protein kinase C C2 domain found
    in all members of the synaptotagmin family"
  misc-feature
    2556..2899
    /gene="syt4"
    /note="C2B repeat region, amino acids 281-394, internal
    repeat homologous to the protein kinase C C2 domain found
    in all members of the synaptotagmin family"
  /citation=[2]
BASE COUNT 1079 a 894 c 986 g 1033 t
ORIGIN
alignment_scores:
  Quality: 1988.00 Length: 425
  Ratio: 4.814 Gaps: 0
  Percent Similarity: 97.176 Percent Identity: 89.412

```

```

alignment_block:
US-09-680-121-2 x MMU10355 ..
Align seg 1/1 to: MMU10355 from: 1 to: 3992
1 MetAlaProIleThrThrSerArgGluGluPheAspGluIleProThrVa 17
1716 ATGGCTCTCTATCACACCACCGCGCTGGAATTGCATGAAATTCACACAGT 1765
17 IValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPhe 34
1766 GGTGGGATCTTCAGTCTTTGGCCGTCCTTCACGCTCTCTCTTTGG 1815
34 IatPrlIeCySGlnArgIysSerSerIysSerAsnIysThrProPro 50
1816 CCTGATCTGCTGTACAGAGAATCAGCCAAATCCAAAGACCTCTCCA 1865
51 TyrLysPheValHisValLeuValGlyValAspIleTyrProGluAsn 67
1866 TACAAAGTTTGTGCACGTCTTAAGAGATTGATATCTAACCCAGAAACT 1915
67 uAsnSerLysLysLysPheGlyAlaAspAspLysAsnGluValLysAsn 84
1916 AAGTAGCCAAAGAAAGTTTGAGAGAGATGACAAAGTCAAGTGAAGGTA 1965
84 ySProAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
1966 AAACCGCTCTCCCAACCTTTCCCTGCATCTTGACCTTGAGAGAACGAGAC 2015
101 LeuAsnGlyAsnPheProLysThrAsnLeuLysProGlySerProSerAs 117
2016 CTCAATGGCACTTCCCAAGCCAAAGCCCAAGCTGACAGCTCTCTTCA 2065
117 pLeuGluAsnAlaThrProLysLeuPheLeuGlyGluLysGluSer 134
2066 TCTGGAATGTCCACCCCAAGCTCTTACGAGACAGAAAAAGAGGCA 2115
134 aIserProGluSerLeuLysSerSerThrSerLeuThrSerGluGluLys 150
2116 ATTCCTCGAGAGCTTGAGATCCAGACTTCCCTCAGCTCAGAGAGAA 2165
151 GlnGluLysLeuGlyThrLeuPhePheSerLeuGluLysThrAsnPheGlu 167
2166 CAAGGAAGAGCTGGGAGACCTCTCTCTCTAGAGTCAACTTCGAGAA 2215
167 gLysAlaPheValValAsnIleLysGluAlaArgGlyLeuProAlaMet 184
2216 GAAAGCATTTCTGTGAACATCAAGGAAGCCAGGCTTACCAGCCATAG 2265
184 sPgIuGlnSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
2266 ACGAGCAATCCATGACCTCTGACCCGTRACATCAAAATGACATCTTACCA 2315
201 GluLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217
2316 GAGAAAGAACACAGAGTGAAGACAGAGTCTCAGAGAGCGTGGACCC 2365
217 AlaPheAspGluThrPheThrPheTyrGlyIleProTyrThrGlnIleG 234
2366 TGTTTTGGTGAACCTTCACATCTATGAAATCCCTTATCCACACATCC 2415
224 InGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
2416 AAGAGCTCTCTGACATTCAGTTCGAGATTTTGGACAGGTTTCTTGA 2465
251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeuSe 267
2466 GATGATGTCTTGGAGAGTCTGTATTCCTTTCAGGAGATTGATATTC 2515
267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284
2516 AGATCGAAAAATGTTAATGACTAGAGATCATCAAGAAATGCTAGAGA 2565

```



122485 GATGAATCCCCACAGTGGGGGATCTTCAGTCATTTGGCTGCTT 122436  
28 ePthVslserLeuphelaTripleCysCysGlnArgLysSerSerLys 45  
122435 CACAGTCTCTCTCTTGGATGAGATGCTGTCTCAGAGAAAAATCATCCAGT 122386  
45 eFasnLysThrProProTyrLysPheValHisValLeuLysGlyValAsp 61  
122385 CTACACAGACTCTCCATACAGTTTGTGATGTGCTTAAGGAGTTGAT 122336  
62 IleTyrProGluAsnLeuAsnSerLysLysPheGlyValAspAspLys 78  
122335 ATTACCTGAGAAACCTAAATACCAAAAGAGTTGGACAGATGATAA 122286  
78 sAnGluValLysAsnLysProAlaValProLysAsnSerLeuHisLeuA 95  
122285 AAATGAAGTAAAGAAATAGCCAGCTGTGCCAAAGAAATTCATTCATCTGG 122236  
95 sPleuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuLys 111  
122235 ATCTTAAAAAGAGATCTCAATGCAATTTTCCCAAAACCACTCATA 122186  
112 ProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuG 128  
122185 CTGGCAGTCTCTGATCTGGAGAAATGCCAACCCGAGCTCTTTTAGA 122136  
128 uGlyGluLysGluSerValSerProGluSerLeuLysSerSerThrSerL 145  
122135 AGGGGAAAAAGATCAGTTTCCCTCAGAGTTTAAAGTCCAGCACTTCCC 122086  
145 euthSerGluGluLysGluLysLeuGlyThrLeuPhePheSerLeu 161  
122085 TTACTTCAGAGAGAAACAAGAGAGCTGGAGACTCTCTTCTCTCTTA 122036  
162 GluTyrAsnPheGluArgLysAlaPheValValAsnIleLysGluAlaAr 178  
122035 GAATACAACTTCGAGAGAAAAAGCATTTGTGTCATATATCAAGAGAGCCG 121986  
178 gGlyLeuProAlaMetAspGluInSerMetThrSerAspProTyrIleL 195  
121985 TGGCTTGGCCAGCATGATGAGAGCATGAGACCTTGACCCCATATATCA 121936  
195 ySmethrIleLeuProGluLysLysHisLysValLysThrArgValLeu 211  
121935 AAATGACGATCTCCAGAGAGAGAGCATTAAGTGAAACTAGAGTCTG 121886  
212 ArgLysThrLeuAspProAlaPheAspGluThrPheThrPheTyrGlyI 228  
121885 AGAAAAACCTTGGATCCAGCTTTTGATGAGACCTTACATTTATGGGAT 121836  
228 eProTyrThrGlnIleGlnGluLeuAlaLeuHisPheThrIleLeuSer 245  
121835 ACCCTACACCCCAATCCAAAGAAATGGCTTGACCTTCACAAATTTTGAGT 121786  
245 heAspArgPheSerArgAspAspIleIleGlyGluValLeuIleProLeu 261  
121785 TTACACAGTTTTCAGAGATGATATCATTTGGGSAAGTTCTAATTCCTCTC 121736  
262 SerGlyIleGluLeuSerGluGlyLysMetLeuMetAsnArgLysIleI 278  
121735 TCGGGAATTAATTAATCTGAAGAAAAATGTAAATGATAGAGATCAT 121686  
278 eLysArgAsnValArg 283  
121685 CAAGAGAAATGTTAGGTAATTTATTTTCAGTGTAAAGTATTTATTA 121636  
283 283  
121635 CAATCTTTATGAGCATTTTGGTGTCTAAGGAATTTTATTAATATGAT 121586  
283 283

121585 AACCTGTACATATCATCCTTACCAGTGTAAATTAATTAATTAATATC 121536  
283 283  
121535 TAATGAGATCAAGGAGATAGAAACAGATGAATAATTAATTAATCATAC 121486  
283 283  
121485 TGAATACTCAAGTGTTTAATGATGTAAATAATATAGCATTTCTTAT 121436  
283 283  
121435 GGATATTTAGATATCACCTATATAATACTTTTGTAAATCTATCATAA 121386  
283 283  
121385 ATATTTTGACAGACACTGCTACTTTAAATACAGGTGAATATTTCTTAT 121336  
283 283  
121335 CATACGAAGTTAACAAGAGATTTCAAGGTCAACATTAATGAATTTAA 121286  
283 283  
121285 ATTGAATTTAAATTAATCAATTTCTTCTGAAGATTTGGAATTTGAT 121236  
283 283  
121235 TGACTTTCCATGGAAATTCCTTCAAGGAAGGCATACGTCATTTTCA 121186  
283 283  
121185 ATTATAATGCTCATCTCTAATTTGTATAGTAGAAATGAGTCCGCTAT 121136  
283 283  
121135 TTTTGTAGTATACTACTCTGTAGTCTGGGTTTATTAGAAAATAGTAA 121086  
283 283  
121085 GACATCAATGAAACAAATTTAAACCCATTTTCACTTAAATTTGAAACATC 121036  
283 283  
121035 GATTTCAACAGTGTATATCAGATAACTTCTCTATATATACATGAT 120986  
283 283  
120985 TATATAAATCATATAGACATCTAATTTGTGTTTTAAACAACATAGTGT 120936  
283 283  
120935 TTCAGAGCTGAAAAACCTGAGAGCTATTTGACTGTCTCTCTATTTG 120886  
283 283  
120885 AAATTAATGACGTCAGAGTTTCAGATACATTTCCAGACCTGTTTGAGACA 120836  
283 283  
120835 CCCGGGTGTTAAGTATACAGTGGGAATAAGATCTTCTCTACTCTT 120786  
283 283  
120785 ATCTCTCTAGAGTGCCTACTATACAGCTGTGCCATATTAATTAATTTTGT 120736  
283 283  
120735 TAAATATATTCAGATTTAGTTTGTGTCAAGCATCATCTCAGTTGCAA 120686  
283 283  
120685 ACAACACCAAGTAAATATATGTTACTTTCAGTCAAACTTTACTTTG 120636

283	.....	283	.....	323	.....	323	.....
120635	AATTATTGTAATCTTATCTATTTGCTTTAAAGTACAAATGAATCTAAG	120586	.....	119735	CACCTACTCTTACTCTCTCTATCTTTCTTCTAATACCTTATGTATCTGCTA	119688	.....
283	.....	283	.....	323	.....	323	.....
120585	TAGCAGCCAAAGAAGACGCTATGACACTGTCCAAAAACAATGCTCTCCA	120536	.....	119685	TAAAGTTAGCTTTTGTATGTATGACTAGCTTACAAAAATGAACAAA	119636	.....
283	.....	283	.....	323	.....	323	.....
120535	ATGCTTAAGATCAATTCATCTCTGCTTAATGAACCTTGCTGTGTG	120486	.....	119635	ACGCCAATTTTATGTTGTATGACAAATGAAACAAAATAAGTACAGT	119586	.....
283	.....	283	.....	323	.....	323	.....
120485	AAAGAGCAGAGATTTTAACTCATTAATCCCTGCGCAGAGATGTAG	120436	.....	119585	TTAAATTTACCTGATTTAAATCTTAACCTTACCTTCAATGCTGCTTAA	119536	.....
283	.....	283	.....	323	.....	323	.....
120435	TTGGGATCAGATTAATTTCTCACTGGGTGTAGTTCAAATTTCAATGTGC	120386	.....	119535	TGATTTGGTATGTTTCTTAACTGTACAACTGGACATATGATCTCAT	119486	.....
283	.....	283	.....	323	.....	323	.....
120385	GGTCTAAAGATTTTGAATCAATCAACTTATCTAATTTTGTGATGA	120336	.....	119485	CCTGAGGGCCTTGTGACACTGAAATACATCTTTCAGATATGATC	119436	.....
283	.....	283	.....	323	.....	323	.....
120335	GTATGTGATCAATTTGACCCACACATCTCATGCAAGATTTAGATAA	120286	.....	119435	TAAATATATTATTATTATTAAATGATGGAGTTGAACAATGTCTTAA	119386	.....
283	.....	283	.....	323	.....	323	.....
120285	GCTTATTTACTATTATTAATCTTTGAAAAAGTAGAGATGTTATTTAAACCA	120236	.....	119385	AGATTAGATTGGAAGCTAAATTAATCAATATGTCACACAGAAAAAG	119336	.....
283	.....	283	.....	323	.....	323	.....
120235	AAAGCCAAATTAATCTGCTGTAGAGATTAATTTAAATATCTTTAAACA	120186	.....	119335	GCMAAGTAATTGTGATTAAGTAGCATGAACCTTAAATCTCATATCTAAC	119286	.....
283	.....	283	.....	323	.....	323	.....
120185	TTTCTAGTAAATGGGTGTGCCAATTTATTCCTAATTTTAACTCTCTGTA	120136	.....	119285	AATGACACAGAGAGCTCAACACTAGAGAAACGAAACGAATCCAGCTG	119236	.....
283	.....	283	.....	323	.....	323	.....
120135	GTACCAAAACAGACCAAGATGGAGGAGGTGTTTCAACAGCAGAT	120086	.....	119235	TGACTTAAGAAATTCATRAAAGTTAAACTAAAAAGAAATTTGCCATTGT	119186	.....
283	.....	283	.....	323	.....	323	.....
120085	GGTACTGTCTCAAAAGCTTTAGTACTCTTCAATCTTGAATATATATA	120036	.....	119185	GGACTCTAAATCTCATTCATTCGATGAGAAAAACAGCCAAAAAGTCAGGAAAT	119136	.....
283	.....	283	.....	323	.....	323	.....
120035	ATTAATAATAGTAAGATTAATTAATTAATGACAGATGTTAATTAACCTTAA	119986	.....	119135	AGACTCACCCAAAGTCATACAGTCACTAAGAGCTACATGTTGCTAATAT	119086	.....
283	.....	283	.....	323	.....	323	.....
119985	TGAAGTACACACTTCAAGTTAAAAAGATAGAATACCTTATATCATAT	119936	.....	119085	ATTTTTCCTTCTATCTCATTTCTCTGAGTAAGTTAATATATACAGAA	119036	.....
284	.....	284	.....	323	.....	323	.....
119935	TCTCTTTCTGCAAGAGTCTTCAGACGGGGTGAATTAATGATCTCTCTCT	119886	.....	119035	TGGGTTCATAGACATTTCTATTTCTTCTCTTGTGTACTCTTTCC	118986	.....
296	yslyrGlnserThrIAsnthrLeuThrValValIleLeuysAlaIArg	312	.....	323	.....	323	.....
119885	GCTATCAGTCCACACCAAAACACTTAACGTGTGTCTTAAAAACCTCGA	119836	.....	118985	TATCATTAATTTTTCCTTATTTCTATTTAGAACATTTTCTCTTAT	118936	.....
313	HisLeuProLysSerAspValSerGlyLeuSer	323	.....	323	.....	323	.....
119835	CATCTCCCTTAATCTGATGTCTCCGACCTTTCAGAGTAAGATGCTTCTGA	119786	.....	118935	TTTCCATATTCTTCTTATTTTCAATTCGTGTGACAGAGAAATGCCAAG	118886	.....
323	.....	323	.....	323	.....	323	.....
119785	AATATTCACATTCACCTGTATGATTGGAAATGTAATTTGACTGCAATGAGCC	119736	.....	118885	TTTATTCATATTTTCTGTTGTTTGAAGTAATTTTCTGTGTAACCAAT	118836	.....
				323	.....	323	.....



```

* 129423 129522: gap of 100 bp
* 129523 138872: contig of 9350 bp in length
* 138873 138972: gap of 100 bp
* 138973 145365: contig of 6393 bp in length
* 145366 145465: gap of 100 bp
* 145466 151207: contig of 5742 bp in length
* 151208 151307: gap of 100 bp
* 151308 155673: contig of 4366 bp in length
* 155674 155773: gap of 100 bp
* 155774 160475: contig of 4702 bp in length
* 160476 160575: gap of 100 bp
* 160576 165680: contig of 5105 bp in length
* 165681 165780: gap of 100 bp
* 165781 169630: contig of 3850 bp in length
* 169631 169730: gap of 100 bp
* 169731 174769: contig of 5039 bp in length
* 174770 174869: gap of 100 bp
* 174870 178072: contig of 3203 bp in length
* 178073 178172: gap of 100 bp
* 178173 180351: contig of 2179 bp in length
* 180352 180451: gap of 100 bp
* 180452 181295: contig of 844 bp in length
* 181296 181395: gap of 100 bp
* 181396 184552: contig of 3157 bp in length
* 184553 184652: gap of 100 bp
* 184653 186282: contig of 1630 bp in length
* 186283 186382: gap of 100 bp
* 186383 188196: contig of 1814 bp in length
* 188197 188296: gap of 100 bp
* 188297 189386: contig of 1090 bp in length
* 189387 189486: gap of 100 bp
* 189487 190550: contig of 1064 bp in length.

```

# FEATURES

```

SOURCE
1..190550
/organism="Homo sapiens"
/dbxref="taxon:9606"
/chromosome="18"
/map="18p11.2"
/clone="RP11-748120"
1..16718
/misc_feature
/feature="assembly_fragment"
16719..30449
/misc_feature
/feature="assembly_fragment"
30500..45607
/misc_feature
/feature="assembly_fragment"
45708..59752
/misc_feature
/feature="assembly_fragment"
59853..75894
/misc_feature
/feature="assembly_fragment"
75995..87034
/misc_feature
/feature="assembly_fragment"
87135..97062
/misc_feature
/feature="assembly_fragment"
97163..106435
/misc_feature
/feature="assembly_fragment"
106536..114800
/misc_feature
/feature="assembly_fragment"
114901..121600
/misc_feature
/feature="assembly_fragment"
121701..129422
/misc_feature
/feature="assembly_fragment"
129523..138872
/misc_feature
/feature="assembly_fragment"
138973..145365
/misc_feature
/feature="assembly_fragment"
145466..151207
/misc_feature
/feature="assembly_fragment"
151308..155673
/misc_feature
/feature="assembly_fragment"
155774..160475
/misc_feature
/feature="assembly_fragment"
160576..165680
/misc_feature
/feature="assembly_fragment"

```

```

misc_feature 165781..169630
/feature="assembly_fragment"
misc_feature 169731..174769
/feature="assembly_fragment"
misc_feature 174870..178072
/feature="assembly_fragment"
misc_feature 178173..180351
/feature="assembly_fragment"
misc_feature 180452..181295
/feature="assembly_fragment"
misc_feature 181396..184552
/feature="assembly_fragment"
misc_feature 184653..186282
/feature="assembly_fragment"
misc_feature 186383..188196
/feature="assembly_fragment"
misc_feature 188297..189386
/feature="assembly_fragment"
misc_feature 189487..190550
/feature="assembly_fragment"

```

```

alignment_scores:
Quality: 1624.00 Length: 1352
Ratio: 3.923 Gaps: 4
Percent Similarity: 30.621 Percent Identity: 30.473

```

alignment\_block:

US-09-680-121-2 x AP001797 ..

Align seg 1/1 to: AP001797 from: 1 to: 190550

```

12 AspGluIleProThrValIcylIlePheSerAlaPheGlyLeuValph 28
9356 GATGAATATCCACACAGTGGGATCTTCAGTTCATTCATTCGCTGCTT 9405
28 ethValSerLeuPheAlaTrpIleCysGlyIleArgLysSerLeu 45
9406 CACAGTCTCTCTTGCATGATCTGCTGCAAGAAATCATTCACAGT 9455
45 eRanLysThrProProlYrLysPheValHisValLeuLysGlyVala 61
9456 CTACACAGTCTCTCTTGCATGATCTGCTGCAAGAAATCATTCACAGT 9505
62 IleYrProGluAsnLeuAsnSerLysLysPheGlyAlaAspAsp 78
9506 ATTACCTCAAAACCTAATAGCAAAAAGAGTTGAGCAGATGATA 9555
78 sAsnGluValLysAsnLysProAlaValProLysAsnSerLeuHisLeu 95
9556 AATGAAGTAAAGATAGACGACGCTGCTGCCAAAGATTCATTCATCTG 9605
95 spleuGluLysArgAspLeuAsnGlyAsnPheProlYrLysLeuLys 111
9606 ATCTTGAAGAGAGATCTCAATGCAATTTTCCAAACCAACCTCAAA 9655
112 ProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeu 128
9656 CCTGCGAGTCTCTCTGATCTGAGATGCAACCCGAGCTCTTTAGA 9705
128 uGlyGluLysGluSerValSerProGluSerLeuLysSerSerThrSer 145
9706 AGGGGAAAAGAGTCACTTCCCTGAGAGTTTAAAGCCAGCCTTCCC 9755
145 eutHisSerGluGluLysGlnGluLysLeuGlyThrLeuPhePheSerLeu 161
9756 TTACTTCGAGAGAGAAACAAGAGAGCTGGAACCTCTCTCTCTCTTA 9805
162 GluYrAsnPheGluArgLysAlaPheValAlaAsnIleLysGluAla 178
9806 GATATCAACTTCAGAGAAAGCATTTGTGTCATATCAAGAGAGCCG 9855
178 gGlyLeuProAlaMetAspGluGlnSerMetThrSerAspProYrIle 195
9856 TGGCTGCGACCATGATGAGCATGATGATGATGATGATGATGATGAT 9905

```



195 ysmethrilleuProglulylslyshislyvalylstrarqValleu 211  
|||||  
9906 AATGACGATCCTCCAGAGAAAGACATAAAGTGAAGAGAGAGCTG 9955  
212 ArglystrhLeuaspProalaphaspGluThrPheThrPheTyrGlyI 228  
|||||  
9956 AGAAAACTGGATCCAGCTTTGTAGACCTTACATTTATGGGAT 10005  
228 eProtyrThrGlnIleGlnGluLeuAlaLeuHisPheThrIleLeuSerP 245  
|||||  
10006 ACCCTACACCCAAATCCAAAGAAATGGCTTGCACTTCACAAATTTGAGTT 10055  
245 heaspArgPheSerArgaspPaspIleIleGlyGluValLeuIleProLeu 261  
|||||  
10056 TTCACAGGTTTTCAGAGATGATGATCATTTGGGGAATCTCAATTCCTC 10105  
262 SerGlyIleGluLeuSerGluGlyLysMetLeuMetAsnArgGluIleI 278  
|||||  
10106 TCGGGAATTTGAATTTATCTGAAGGAAAAATGTTAATGAATAGACATCAT 10155  
278 elysArgAsnValArg..... 283  
10156 CAAGAGAAATGTAGGGTAATTTATTTTCAGTGCCTTAAAGTATTTATAAA 10205  
283 ..... 283  
10206 CATCTTTATGAGCATTTTGGTGTCTAAGAAATTTTATTAATATGAT 10255  
283 ..... 283  
10256 AAGCTGTACATATCATCTTACCATGTTTAAATTATTAATTAATATATC 10305  
283 ..... 283  
10306 TAATGAGATCAAGGAATAGAACAGATGAATAAAAATTAATCCATAC 10355  
283 ..... 283  
10356 TGAATACTCAAGGTGTTTAAATGATGTAAATAATTAAGCATTTCTTAT 10405  
283 ..... 283  
10406 GGATATTTAGATATCATCTATATAACTTTTGTATCTCTATCATATA 10455  
283 ..... 283  
10456 ATATTTTGACAGACACTGCTACTTAAAAATACAGTGGAATATTTCTTAT 10505  
283 ..... 283  
10506 CATTAACAGTTAACAAAGAGATTTCAAGGTCAACATTAATGAATTTAAA 10555  
283 ..... 283  
10556 ATTTGAATTTAAATTAATCAATTCCTCTCTGAAGATTTGGAATTTGAT 10605  
283 ..... 283  
10606 TGACTTTCATGGAATGGCTTTCAAGGAAGAGCATCACGTCATTTTCA 10655  
283 ..... 283  
10656 ATTTAAATGCTCATCTTAATTTGTATAGTAGAAATGAGTCCCGCTAT 10705  
283 ..... 283  
10706 TTTTGTAGTATATCTACTCTGTAGTCTGTGTTTATTAAGAAATAGTAA 10755  
283 ..... 283  
10756 GACATCAATGAAACAATTTAACCCATTTTCACTTAAATTTGAACATC 10805

283 ..... 283  
10806 GATTTCACAGTGTATATACAGATAACTCTCTATATAATACATGAT 10855  
283 ..... 283  
10856 TATAAAAATCAATAGACATCTAATTTGTGTTTTTAAACACATAGTGT 10905  
283 ..... 283  
10906 TTCAGAGCTGAAAAACCTGAGAGCTCATTTTGACTGTCTCTCTATTTG 10955  
283 ..... 283  
10956 AAATTAATGACGGTCAAGGTTAGATATATCCACAGCTTGTGAGACA 11005  
283 ..... 283  
11006 CCGGCTGTGTTAAGTATACAGTGGGAATAGGATCTTCTCTACTCTT 11055  
283 ..... 283  
11056 ATTCTCTTAGAGTGCCTACTTATACAGCTGTCCCATATTAGTTTGTGA 11105  
283 ..... 283  
11106 TAAATATATTGCAATTTAGTTTGTGTCAAGCATCATCTCAGTTGCAA 11155  
283 ..... 283  
11156 ACAACACCAAGTTAATATATTTGTTACTTTCAGTCAAACTTTTACTTTG 11205  
283 ..... 283  
11206 AATTATTGTATCTTATCTTATTTGTCTTAAAGTACAAATGAATCTAAG 11255  
283 ..... 283  
11256 TAGACGCAAAAGAAAGACGCTATGACACATGCCAAAAACAATGCTTCA 11305  
283 ..... 283  
11306 ATGCTTAAGATCAATTCATTCCTTGCTTAATGAAACCTTGCTGTGTG 11355  
283 ..... 283  
11356 AAAGAGCAGAGAAATTTAACTCCATTATCCCTGTGGCCAGCATGTAG 11405  
283 ..... 283  
11406 TTGGGATCAGAAATATCTCAGTGGGTCTGAGTTCAAATTTCAATATGTC 11455  
283 ..... 283  
11456 GGTGCTAAAGAAATTTGAATCAATCAAACTTATCTAATTTTTTGCATGA 11505  
283 ..... 283  
11506 GTATGTAGTCAATTTGACCAACACATCTCATGCAAGAAATAGATTA 11555  
283 ..... 283  
11556 GCTTTATTAATTAATTAATCTTGAAAAAGTAGAGATTTTAAAAACCA 11605  
283 ..... 283  
11606 AAAGCAAAATAATCTCTGTAGATATATTAATTAATCTTTAACA 11655  
283 ..... 283  
11656 TTTCTAAGTAATGGGTGGCCAAATTTATGCAATTTTAACTTCTGTGA 11705  
283 ..... 283

11706 GTACCAACAGACACAGATGCGAGCGATGTTTTCAGACGCGAT 11755  
 283 ..... 283  
 11756 GGTACTGCTCAAAAGCTTTAGTACTCTTCATTCCTTAGAATATTA 11805  
 283 ..... 283  
 11806 ATTAATAATAGTAAGATTATTTAATGACATGTTATTTAACTTTAA 11855  
 283 ..... 283  
 11856 TGAAGTACACACTTCAAAGTTAAAAAGATAGAAATCTTATATCACA 11905  
 284 .....Lys.SerSerGlyArgGlyLeuLeuIleSerLeuC 296  
 11906 TCTCTTCTGCAAGAGCTTCAGAGCGGCGTGTACTGATCTCTCTCT 11955  
 296 YSTYGLINserThrAsnThrLeuThrValValLeuLysAlaArg 312  
 11956 GCTATCAGTCACCAACAACACTTAACCTGGTGCTTTAAAGACTCGA 12005  
 313 HSLeuProLysSerAspValSerGlyLeu.Ser..... 323  
 12006 CATCTGCCATTAATCGATGTCTCGGACTTTTCAGTAAGATGCTTCTGA 12055  
 323 ..... 323  
 12056 AATATTGAGTTCACTGATGATGTAATGATATTGACTGCAATAGAGCC 12105  
 323 ..... 323  
 12106 CACCTACTTCTCTCTCTATCTTTTCTTAATACCTTATGATATGCTA 12155  
 323 ..... 323  
 12156 TAAAGTACGTTTACTTTTATGACTAGCTTTACAAAAAGAAACAA 12205  
 323 ..... 323  
 12206 AGCCCAATTTATGTTGATAGCAATGAGAAACAAATTAAGCTACGT 12255  
 323 ..... 323  
 12256 TTAAATTAAGTGAAGTCAAACTTAACCTTCATTCAGTACGAAAAATA 12305  
 323 ..... 323  
 12306 TGATGTTGATGTTCTTTAACTGATAGCAACGAGATATGCTACTCAT 12355  
 323 ..... 323  
 12356 CTTGAGGGGCTTGTGAGACTAGAAATACATACTTCTGATATATC 12405  
 323 ..... 323  
 12406 TAATTAATTTATTTATTTAATGATGGAGTTGAAACATGGTCTTAA 12455  
 323 ..... 323  
 12456 AGATTGATTTGAGAGCTAAATTAACATATGTCGACAGAAAAAG 12505  
 323 ..... 323  
 12506 GCAAAAGTATTTGTATTAAGATGACATGAAATCTACATATCAAC 12555  
 323 ..... 323  
 12556 AATGACACAGAGAGCTCAACACTAGGAACAGAAAGCAATACACCTG 12605  
 323 ..... 323

12606 TGACTAAGAAATTCATTAAGTTAAAACTAAAAAGATATTCATTTGT 12655  
 323 ..... 323  
 12656 GGACTCTAATCTCATTCATGATGGAACACAGCCAAAGTCAGGGAAT 12705  
 323 ..... 323  
 12706 AGACTACCCAAAGTCATACAGAGCTAAGCTACATGTGTCTAATAT 12755  
 323 ..... 323  
 12756 ATTTTTCCTCTATCTATCTCTGAGTAAGTTAATAGTATCAGAA 12805  
 323 ..... 323  
 12806 TGGGTCCATAGAACATTCATTTCTTCTTCTTGTGCTACTCTTTCC 12855  
 323 ..... 323  
 12856 TATCATTAATTTTTCCTTATTCCTTAATTTAGAACTTTTCTCTTAT 12905  
 323 ..... 323  
 12906 TTTTCATTTCTCTTATTTTCAATTCCTGTGACAGAGAAATGCCAAG 12955  
 323 ..... 323  
 12956 TTTTANCCATATTTTCTGTGTGTGTGTGACTTATTTCTGTGTAACCAT 13005  
 323 ..... 323  
 13006 GTCATTGCGATGCCATGATCTTAAGGSCACTTATCTGATATCAGCT 13055  
 324 .....AspP 325  
 13056 CTCACGCCAGTTATGCCATGTGTAATGCGCTGTATGTTCCACTGCAATC 13105  
 325 TGTValAllyValAsnLeuThrHisAlaLysLysArgLysSerLysLys 341  
 13106 CTTATGTCAAAGTGAAGCTGTACCTGCAAAAGAGATCTCAAGAG 13155  
 342 LysThrHisValLysLysCysThrProAsnAlaValPheAsnGluLeuPhe 358  
 13156 AAGACTATGTGAAGAAATGCAACCCCAATGCACTGTTCAATGAGCTGTT 13205  
 358 eValPheAspLysIleProCysGluGlyLeuGluAspLysSerValGluPheL 375  
 13206 TGTCTTGAATATCTCTTGAGAGGCTTGAAGATATAGTGTGAAATTTT 13255  
 375 euValLeuAspSerGluArgLysSerArgAsnGluValIleGlyGlnLeu 391  
 13256 TGGTTTGGATCTGTGAAGGGGCTCCCAATAGGTAATCGGCGACGTTA 13305  
 392 ValLeuGlyAlaAlaGluGlyThrGlyGlyLysIleStrLysLysLys 408  
 13306 GTCCTTGGTCCAGCAGAGAACTGTGTGAGACACTGGAAGAGAT 13355  
 408 eCysAspLysProArgArgGlnIleAlaLysTrpHisValLeuCysAspG 425  
 13356 CTGTGACTACCCGAGAGACAAATTTGCCAAGTGGCAGCTGTGTGTGATG 13405  
 425 Ly 425  
 13406 GT 13407  
 seq\_name: gb\_hhg:AC087507  
 seq\_documentation\_block: 192022 bp DNA HTG 09-MAY-2001  
 LOCUS AC087507 192022 bp DNA HTG 09-MAY-2001  
 DEFINITION Homo sapiens chromosome 18 clone RP11-748120 map 18, WORKING DRAFT  
 SEQUENCE, 12 unordered pieces.  
 ACCESSION AC087507

```

VERSION      AC087507.3  GI:13940683
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 192022)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE        Homo sapiens chromosome 18, clone RP11-748120
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 192022)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguski,K., Boulikas,A., Brown,A.,
Camargue,J., Campopiano,A., Chao,Y., Collinge,S.,
Collamore,A., Cooke,P., DeArnell,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyns,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagob,B., Heaford,A., Horton,L., Hulme,W., Illiev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Mathews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mienga,Y.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Reta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,
Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Tamas,J., Tesfaye,S., Theodore,J.,
Travis,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zemsek,L., Zimmer,A. and Zody,M.
COMMENT      Direct Submission
              Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On May 4, 2001 this sequence version replaced gi:12957877.
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html

              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu

              ----- Project Information
              Center project name: 748_1-20
              Center clone name: 748_1-20

              ----- Summary Statistics
              Sequencing vector: Plasmid; n/a; 100% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.960731
              Consensus quality: 186183 bases at least Q40
              Consensus quality: 189224 bases at least Q30
              Consensus quality: 190373 bases at least Q20
              Insert size: 194000; agarose-fp
              Insert size: 190922; sum-of-ctnigs
              Quality coverage: 6.1 in Q20 bases; agarose-fp
              Quality coverage: 6.2 in Q20 bases; sum-of-ctnigs

              -----
              * NOTE: This is a 'working draft' sequence. It currently
              * consists of 12 contigs. The true order of the pieces
              * is not known and their order in this sequence record is
              * arbitrary. Gaps between the contigs are represented as
              * runs of N, but the exact sizes of the gaps are unknown.
              * This record will be updated with the finished sequence
              * as soon as it is available and the accession number will
              * be preserved.

              1 6805: contig of 6805 bp in length
              * 6806 6905: gap of 100 bp
              * 6906 8180: contig of 1275 bp in length
              * 8181 8280: gap of 100 bp
              * 8281 9283: contig of 1003 bp in length
              * 9284 9383: gap of 100 bp

```

```

FEATURES
source
* 9384 11168: contig of 1785 bp in length
* 11169 11268: gap of 100 bp
* 11269 13778: contig of 2510 bp in length
* 13779 13878: gap of 100 bp
* 13879 28641: contig of 14763 bp in length
* 28642 28741: gap of 100 bp
* 28742 43445: contig of 14704 bp in length
* 43446 43545: gap of 100 bp
* 43546 62579: contig of 19034 bp in length
* 62580 62679: gap of 100 bp
* 62680 86613: contig of 23934 bp in length
* 86614 86713: gap of 100 bp
* 86714 125447: contig of 38734 bp in length
* 125448 125547: gap of 100 bp
* 125548 157506: contig of 31959 bp in length
* 157507 157606: gap of 100 bp
* 157607 192022: contig of 34416 bp in length.
Location/Qualifiers
1. 192022
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-748120"
/clone_lib="RP11-748120 Human Male BAC"
1. 6805
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
6906. 8180
/note="assembly-fragment"
8281. 9283
/note="assembly-fragment"
9384. 11168
/note="assembly-fragment"
11269. 13778
/note="assembly-fragment"
13879. 28641
/note="assembly-fragment"
28742. 43445
/note="assembly-fragment"
43546. 62579
/note="assembly-fragment"
62680. 86613
/note="assembly-fragment"
86714. 125447
/note="assembly-fragment"
125548. 157506
/note="assembly-fragment"
157607. 192022
/note="assembly-fragment"
clone_end:T7
vector_side:right"
BASE COUNT 60131 a 34828 c 35147 g 60815 t 1101 others
ORIGIN
alignment_scores:
Quality: 1624.00 Length: 1352
Ratio: 3.923 Gaps: 4
Percent Similarity: 30.621 Percent Identity: 30.473
alignment block:
US-09-680-121-2 x AC087507 ..
Align seg 1/1 to: AC087507 from: 1 to: 192022
12 AspguilepProtnrValaGlyIlepheserAlapheglyLeuValph 28
|||||
183467 GATGAATCCCAAGTGGGATCTTCAGTCATTTGGCTGCTT 183516
|||||
28 ethValserLeuPhaIarIleCysGlnArgIysSerSerIys 45
|||||

```

183517 CACAGTCTCTCTTTGCATGATCTGCTCAGAGAAATCATCCAACT 183566  
45 eRAsnlyserProserProlyserPheValHisIleuLysGlyValAsp 61  
183567 CTACACAGCTCTCTCATACAGAGTTTGTCATGCTTAAGGAGTTGAT 183616  
62 lIeTyTProGluAsnLeuAsnSerLysLysPheGlyAlaAspAspLy 78  
183617 ATTTACCTGAAACCTTAATAATAGCAAAAAAGAGTTTGAGCAGATGATPA 183666  
78 sAsnGluValLysAsnLysProAlaValProLysAsnSerLeuHisLeuA 95  
183667 AATATGAAGTAAAGATTAAGCCAGCTGCGCAAGAAATTCATTGCTCTGG 183716  
95 sPLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuLys 111  
183717 ATCTTGAAAAAGAGATCTCAATGCAATTTTCCAAAAACCACTCTCAA 183766  
112 ProGlySerProserAspLeuGluAsnAlaThrProLysLeuPheLeuG 128  
183767 CCTGGCAGCTCTTGATCTGAGATGCAACCCGACGCTCTTTTGA 183816  
128 uGlyGluLysGluSerValSerProGluSerLeuLysSerSerThrSerL 145  
183817 AGGGGAAAAAGAGTCAGTTTCCCTGAGAGTTTAAAGTCCACGACTTCC 183866  
145 euThSerGluGluLysGluLysLeuGlyThrLeuPhePheSerLeu 161  
183867 TTACTTCAGAGAGAGAAACAAGAGAGCTGGGACTCTCTTCTTCCTTA 183916  
162 GluTyTAsnPheGluArgLysAlaPheValAlaAsnIleLysGluAla 178  
183917 GAATACACTTCGAGAGAAAAAGCTTTTGCTAATATACAGAGAACCCG 183966  
178 gGlyLeuProAlaMetAspGluGlnSerMetThrSerAspProLyIleL 195  
183967 TGGCTTGGCAGCGATGATGAGCGAGTGAATGACCTGACCCCATATACA 184016  
195 yseMetThrIleLeuProGluLysLysHisLysValLysThrArgValLeu 211  
184017 AATATGACATCTCCCAAGAGAAAGCATAAAGTAAACTGAGATGCTG 184066  
212 ArgLystrLeuAspProAlaPheAspGluThrPheThrPheTyGlyI 228  
184067 AGAAAAACCTTGATCCAGCTTTTGATGACACCTTACATTCATAGCGAT 184116  
228 eProTyThrGlnIleGlnLysLeuAlaLeuHisPheThrIleLeuSer 245  
184117 ACCCTACACCCCAATCCAGAAATGGCTTGCCTTCACAAATTTTGAGTT 184166  
245 heAspArgPheSerArgAspAspIleIleGlyGluValLeuLeuProLeu 261  
184167 TTGACAGGTTTCAAGAGATGATATCATTTGGGAAGTTCTTAATCTCTC 184216  
262 SerGlyIleGluLeuSerGluGlyLysMetLeuMetAsnArgGluIleI 278  
184217 TCGGGAATTGAAATATCTGAAAGAAAAATGTTATATGATAGAGATCAT 184266  
278 eLysArgAsnValArg 283  
184267 CAAGGAATATGTAGGTAATTAATTTTCAGTGCCTTAAAGTATTTATAAA 184316  
283 283  
184317 CAATCTTTATAGATTTTGGTGTCTAAGAAATTTTATTAATATGAT 184366  
283 283  
184367 AAGCTTGATATATCATCTTACCATGTTTAATTAATTAATATAATC 184416  
283 283  
184417 TAATGAGATCAAGGAAATGAAACAGATGAATTAATTAATCATTAAC 184466

283 283  
184467 TGAATATCTCAAGTGTTTTAATGATGTAATAATAATAGCATTTCTTAAT 184516  
283 283  
184517 GGATATTTAGATATCATCCTATATTAATACTTTTGTATATCTATGATAA 184566  
283 283  
184567 ATATTTTGACAGACACTGCTACTTTAAATACAGTGAAATATTTCTTAAT 184616  
283 283  
184617 CATTAAGCAAGTTAACAAAGAGATTTCAAGTCAACATTAATGAATTTAAA 184666  
283 283  
184667 ATTGAATTTAAATTAATTAATCAATTTCTTCTGAAGATTTGGAATTTGAT 184716  
283 283  
184717 TGACTTCCATGGAATTTGCTTCAAGAGAAAGGATCAGCTCATTTTCA 184766  
283 283  
184767 ATTTAATAGCTCATCTCTAATTTGTATAGTAAGAAATGAGTCCCGCTAT 184816  
283 283  
184817 TTTTGTATATATCTACTCTGCTGCTGCTGCTGCTTATTAAGAAATAGTAA 184866  
283 283  
184867 GACATCAATGAACAATTTAACCCCATTTTCACTTAAATTTGAAACTC 184916  
283 283  
184917 GATTTCAACAGTGATATACGAATTAATCTCTCTATTAATACAAATGAT 184966  
283 283  
184967 TATTAATAATCAATAGACATATATTTGTTTTTAAACAAATAGTGT 185016  
283 283  
185017 TTCAGAGCTGAACACCTGAGAGCTCATTTTGACTGTCTCTCTCTATTG 185066  
283 283  
185067 AATTTAATGACGTCAGGTCAGTTCAGATPATCTCCACAGCTTGTGAGACA 185116  
283 283  
185117 CCCGCTGGTTAAGTATACATGAGTGGGAATTAAGATCTTTCTTCTACTCT 185166  
283 283  
185167 ATTCTCTTAGAGTGCTACTTATATACAGCTGTCCCATATTAATTTTGTGA 185216  
283 283  
185217 TAAATATATGCGATTTAGTTTGTGTGACGCCATCAGCTCAGTTGCAA 185266  
283 283  
185267 ACAACACCAAGTTAATATATTTGTTACTTTCAGTCAAACTTTACTTGG 185316  
283 283  
185317 AATTAATGATATCTTATCTATGTGCTTTAAAGTACAAATGAATCTTAAG 185366

283 ..... 283  
185367 TAGACGCAAGAGAGCGCTATGACACTGTCCAAAAACATGCTCTCA 185416  
283 ..... 283  
185417 ATGCTTAAGATCAATTCATTCCTGGCTTAATGAAACTGTCTGTGTG 185466  
283 ..... 283  
185467 AAGAGGAGAGAAATTTAACTCCATATATCCCTGTGGCCAGCAATGTAG 185516  
283 ..... 283  
185517 TTGGGATCAGAAATATTCCTACTGGGTCTGAGTTCAATTTCAATATGTC 185566  
283 ..... 283  
185567 GGTGCTAAAGAAATTTGAAATCAAACTATATTAATTTTTCATGA 185616  
283 ..... 283  
185617 GTATGTGTAAGTATTTGACCCACACATCTCATGCAAGAAATTAGATAA 185666  
283 ..... 283  
185667 GCCTTATTACTATTAACTTTGAAGAGTAGAGATGTATTAAAAACA 185716  
283 ..... 283  
185717 AAAGCCAAATAATCCTGTGTAGATATATTAATTAATCCTTTAACA 185766  
283 ..... 283  
185767 TTCTTAAGTAAGGTGTGGCCATTTATTTGCCATTTTAACCTTCGTA 185816  
283 ..... 283  
185817 GTACCAACAAGACACAGATGCGCAGCAGTGTTTTCAAGCAGCAT 185866  
283 ..... 283  
185867 GGTACTGCTCAAAAGCTTTAGTTACTCTCAATTCCTTAGAATATTTAA 185916  
283 ..... 283  
185917 ATAAAAATAAGTAGATTATTAATTAATGACGATGTTAATTAACCTTAA 185966  
283 ..... 283  
185967 TGAAGTACACACTTCAAGTTAAAAAGATAGAATACTTATATACATT 186016  
284 ..... 284  
186017 TCCTCTTCGACAGAGCTTCAGGAGCGGGTGAAGTACTGATCTCTCT 186066  
284 ..... 284  
284 ystyrglnserthrthrasnthrleuthrvalvalleulslalarg 312  
186067 GCTATCAGTCCACCAACAACCTTAAGTGTGTCTTAAAAAGCTCA 186116  
313 Hisleuprolysseraspvalserglyleu.Ser..... 323  
186117 CATCTGCTTAATCTGATGTGTCTCGGACTTTCAGGTAAAGATGCTTCTGA 186166  
323 ..... 323  
186167 AATATTCAGATTCAGTATGATGGAATGATTTGACTGCATGAGCC 186216  
323 ..... 323  
186217 CACCTACTTACTCTTCTATCTTTTCTTAATACCTTAATGATATCTGCTA 186266  
323 ..... 323

186267 TAAAGTTAGCTTTTACTTTATGACTAGCTTTACAAAAATGAACAAA 186316  
323 ..... 323  
186317 ACGCCAAATTTTANGTGTATAGCACAATGAGAAACAAATAAGCTACAGT 186366  
323 ..... 323  
186367 TTAATAATTACCTGAGTTCAAATCCTAACCCCTTCATTCAGTACGAATA 186416  
323 ..... 323  
186417 TGATTGGTATGTTCTTAACCTGACAACTGACATAATAGTACTCAT 186466  
323 ..... 323  
186467 CCTGAGGGGCCCTGTGAACTAGAAATACACTCTCTGATATGATC 186516  
323 ..... 323  
186517 TAATAATATTAATTTATTAATGATGGAGTGAACAACATGCTTAAA 186566  
323 ..... 323  
186567 AGATTAGATTGTGAAGCTAAATAATTACCATATGTGTGCCACAGAAAAAG 186616  
323 ..... 323  
186617 GCAAAAGTAATTTGTAAAGTAGACATGAACTTAAATCTACATCTAAC 186666  
323 ..... 323  
186667 AATGACACAGAAAGCTCACACACTAGAACAGAAAGCAATCCAGCCTG 186716  
323 ..... 323  
186717 TGACTAAAGAAATTCATTAAGTTAAAGTAAAGATATTTGCCATTGT 186766  
323 ..... 323  
186767 GGAAGTAAATCTCATTCGATGGGAAAAACAGCCAAAGTCAAGGAAT 186816  
323 ..... 323  
186817 AGACTCACCAAGTACATACAGTCAAGTAAAGTACATGTTGTCTAATAT 186866  
323 ..... 323  
186867 ATTTTTCCTTATCTTATCTCTGTGAGTAAGTTAATTAATCAAGAA 186916  
323 ..... 323  
186917 TGGGTCCATAGAACATCTATTTCTTCTTCTTGTGTACTCTTTCC 186966  
323 ..... 323  
186967 TATCATTAATTTTCCCTATTCCTAATTTAGAACTTTTCTCTTAT 187016  
323 ..... 323  
187017 TTTCCATATTCCTCTTATTTTCAATTCCTGTTGACAAAGAAATGCCAAG 187066  
323 ..... 323  
187067 TTTATCCAAATTTTCTGTTGTTTGTGACTAATTTCTCGSTAACCAAT 187116  
323 ..... 323  
187117 GTCAATGGCATGCCATGATCTTAAGGCACTGTTATCTGATATCAAGT 187166  
324 ..... 324  
asp 325  
|||

```

187167 CTCGAGCCAGTATGATCCATTGTAATGCTGTATGTTCCACTGCAGATC 187216
325 rctyVallyValasnlLeuTyrlHisAlaLysArgIleSerLysLys 341
187217 CCTATGCTCAAGTGAACCTGTACCATGCTCAAAAAGAGAAATCTCCAAAGAG 187266
342 LysThrHisValLysLysCysThrProAsnAlaValPheAsnGluLeuPhe 358
187267 AAGACATCATGTCAAGAAATGACACCCCAATGATGCTTCATGATGCTGTT 187316
358 eValPheAsnIleProCysGluGlyLeuGluAspIleSerValGluPheL 375
187317 TGTGCTTGTATCTCTTGTGTAGGGCTTGAAGATATAGTGTGAATTTT 187366
375 euValLeuAspSerGluArgLysSerArgAsnGluValIleGlyGlnLeu 391
187367 TCGTTTGTGATTCGTAAGAGGGGCTCCGAATATGATTCGGCAGTTA 187416
392 ValLeuGlyAlaAlaAlaGluGlyThrGlyGlyGluHisTrpLysGluI 408
187417 GCTTGGGTGCGACGACGAGAGACTGTGTGAGAGCACTGGAAAGAGAT 187466
408 eCysAspTrpProArgArgGlnIleAlaLysTrpHisValLeuCysAspG 425
187467 GGTGTACATCACCCAGAGAGACAAATGTCGAAGTGGACGTCGTGTGATG 187516
425 ly 425
187517 GT 187518

seq_name: gb.htg:AC091039
seq_documentation_block:
LOCUS AC091039 159135 bp DNA HTG 11-AUG-2001
DEFINITION Homo sapiens chromosome 18 clone RP11-403A5 map 18, *** SEQUENCING
IN PROGRESS ***, 3 unordered pieces.
ACCESSION AC091039.3 GI:15148197
VERSION AC091039.3
KEYWORDS HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (base 1 to 159135)
Birten,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-403A5
Unpublished
2 (bases 1 to 159135)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barnen,N., Bastien,V., Boguslavsky,L., Bouhassalier,B., Brown,A.,
Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., DeArrellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,M., Gage,D.,
Gajagan,J., Gardyna,S., Glende,S., Goylete,M., Graham,L.,
Grand-pierre,N., Hagos,B., Heatford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., Larocque,R.,
Lamares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Marcquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Menais,L.,
Mihova,T., Mieng,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollard,V., Raymond,C.,
Rella,R., Rieback,M., Riley,R., Rise,C., Rogov,R., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,
Severly,P., Sougnuez,C., Spencer,B., Stange-Thamann,N.,
Sodjanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zaloun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 11, 2001 this sequence version replaced gi:14316523.

```

```

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13167
Center clone name: 403_A5
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 69068: contig of 69068 bp in length
* 69069 69168: gap of 100 bp
* 69169 73453: contig of 4285 bp in length
* 73454 73553: gap of 100 bp
* 73554 159135: contig of 85582 bp in length.
Location/Qualifiers
1. 159135
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-403A5"
/clone_id="RP11-403A5"
Human Male BAC"
BAC COUNT 49874 a 28773 c 28100 g 52128 t 260 others
ORIGIN
alignment_scores:
Quality: 1603.00 Length: 1353
Ratio: 3.881 Gaps: 5
Percent Similarity: 30.525 Percent Identity: 30.451
alignment_block:
US-09-680-121-2 x AC091039/rev ..
Align seg 1/1 to reverse of: AC091039 from: 1 to: 159135
12 AspGluIleProThrValVal.GlyIlePheSerAlaPheGlyLeuValP 28
150797 GATCAAAATCCCCACAGCTGTCGATGTCAGTCATTGGCGCTGTC 150748
28 heThrValSerLeuPheAlaTrpIleCysCysGlnAArgLysSerSerLys 44
150747 TCACAGTCTCTCTTTCATGAGATCTCTCTCAGAGAAATATATCCAA 150698
45 SerAsnLysThrProProTrpTrpLysPheValHisValLeuLysGlyVal 61
150697 TCTTACAGACTCTCTCCATACAAAGTTTGTGATGCTTTAAGGAGTTGA 150648
61 PlietyrProGluAsnLeuAsnSerLysLysLysPheGlyAlaAspAspL 78
150647 TATTTTACCTCGAAACCTTAATATGCAAAAGATTTTGAGACAGATGATA 150598
78 ysaAngluValLysAsnLysProAlaValProLysAsnSerLeuHisLeu 94
150597 AAAATGAGATGAAGATTAAGCCAGCTGTGCCAAAGAAATTCATTGCA 150548
95 AspLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuTy 111
150547 GATCTTGAAAGAGAGATCTCAATGCAATTTTCCCAAAACCAACTCAA 150498
111 sProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuG 128

```

150497 ACCGGCAGCTCTTGTGANTCTGGAGATGCACCCGAAAGCTCTTTTAG 150448  
128 LUGLYGUGYSGUSERVALSERPROGLUSERLEUYSSERSETHR 144  
150447 AAGGGGAAAAAGTGCAAGTTCCCTGAGAGTTTAAAGTCACACCTTCC 150398  
145 LEUTHSERGLUGLYSGINGLYLSLEUGLYTHRLEUPHESERLE 161  
150397 CTTACTTCAGAGAGAGAAACAGAGAACTGGAGACTCTCTTCTTCCTT 150348  
161 UGLUTYRASNPHGLUARGLYSALAPHEVALYALASNILELYSGLUALA 178  
150347 AGAATACAACTTCGAGAGAAAACATTTGGTCAATATCAAGAGAGCCC 150298  
178 RGLYLEUPROALAMETASPGUGLNSERMETHRSERASPROTYRILE 194  
150297 GTGGCTTGGCCAGCATGAGAGACACGTGACCTGACCATATATC 150248  
195 LYSMETTHRILEUPROGLUTLYSLYSHISLYSVALYSTRARGYALLE 211  
150247 AAAATGACGATCTCCCAAGAGAAAGCATAAAGTGAAGCTGAGTGTCT 150198  
211 UARGLYSTHRLEUASPPROALAPHEASPGIUTHRPHETHRPHETRYC 228  
150197 GAGAAAAACCTTGATGACGCTTTGATGAGACCTTACATCTCATGGGA 150148  
228 LEPROTYRTHRGINILEGLINLULEUALLEUHSIPHERHILEUSER 244  
150147 TACCCCTACACCACAAATCCAGAAATGGCTTGACATTCACAAATTTGAGT 150098  
245 PHEASPARPHESEARGPSASPILEILEGLYUVALLEULEPROLE 261  
150097 TTTCAGACAGTTTCAAGAGATGATCATTTGGGAGTTCTAATTCCTCT 150048  
261 USERGLYLEGLULEUSERGLUGLYLSMETLEUMETASNARGLUILET 278  
150047 CTCGGGAATTTGAATTAATCTGAAGAAAAATGTAATGAATGAGAGATCA 149998  
278 LELYSATGPSNVALARG..... 283  
149997 TCAGAGAAATGTAGGTAATTTATTTTCAGTGTAAAGTATTTATAA 149948  
283 ..... 283  
149947 ACAATCTTTATGTAGCATTTTGGTGTCTAAGGAATTTTATTAAATGA 149898  
283 ..... 283  
148997 TAACTGTGATCATATCATCTTACATGTTTAAATTAATTAATTAAT 149848  
283 ..... 283  
149847 CTATATGAGATCAAGGAATGAGACAGATGAATAAATTAATTAATCA 149798  
283 ..... 283  
149797 CTGAATACTCAAAAGTGTTTAATGATGTAATAAATATAGCATTTCT 149748  
283 ..... 283  
149747 TGGATATTTAGATATCATCCTATATAATACCTTTGTAAATCTCATATA 149698  
283 ..... 283  
149697 AATATTTTGACACACACTGCTACTTTAAATACAGTGAATTTTCTTAT 149648  
283 ..... 283  
149647 TCATAAGCAAGTTAACAAAGAGATTTCAAGTCAACATAATTTGAATTA 149598  
283 ..... 283  
149597 AATGAAATTTAAATTAATCAATCTCTCGAGATTGGAAATTTGTA 149548

283 ..... 283  
149547 TTGACTTTCATGGAATTTGCTTTCAAGAGAGGATCATGCAATTTTC 149498  
283 ..... 283  
149497 AATTTAATGTCATCTCTAATTTGATAGTAGGAATGATGCCGCTA 149448  
283 ..... 283  
149447 TTTTGTAGTTACACTACTGTGTAGTTCTGGGTTTATGAAAAATAGTA 149398  
283 ..... 283  
149397 AGACATCAATGAACAATTTAACCCCATTTTCACTTTAAATTTGAAACT 149348  
283 ..... 283  
149347 CGATTTCAACAGTGTATATCAGAAATACCTTCCTCTATATCAATGAT 149298  
283 ..... 283  
149297 TTATAAAAATCAATAGACATATAATTTGTGTTTTTAAACAATAGTG 149248  
283 ..... 283  
149247 TTTTCAGACTGAAAACCTGAGAGCTCATTTTGACTGTCTCTCTCTATT 149198  
283 ..... 283  
149197 GAATTTATGACGGTCAAGTTTCAGATACATTCACAGCTGTTTGAGAC 149148  
283 ..... 283  
149147 ACCGGCTGTTAAGTATACAGTGGGAATAGATCTTCTCTACTCT 149098  
283 ..... 283  
149097 TATTCCTTAGAGTGCTTACTTATACAGCTGCCATATTAGTTTTGT 149048  
283 ..... 283  
149047 ATAAATATATTGCAGTTTATGTTTTGTGCAGCATCTCTGAGTTGA 148998  
283 ..... 283  
148997 AACAAACACCAAGTTAATATATGTGTACTTTCAGTCAAAACTTTACTT 148948  
283 ..... 283  
148947 GAATTTATGATATCTTATCTATTGTCTTTAAAGTACAAATGAAATCTAA 148898  
283 ..... 283  
148897 GTAGACAGCAAGAAGAGACGCTATGACACTGCCAAAAAANATGCTCTCC 148848  
283 ..... 283  
148847 AATGCTTAAGATCAATTCATTCCTTGGCTTAATGAAACTTGCTGTGT 148798  
283 ..... 283  
148797 GAAAGAGCAGAGATTTTAACCTCATATTATCCCTGTGGCCAGGAATGTA 148748  
283 ..... 283  
148747 GTTGGATCAGAAATATTTCTACTGGCTGTGAGTTCAAAATTTCAATATGT 148698  
283 ..... 283  
148697 CGGTCTAAAGAAATTTGAAATCAAACTATCTAATTTTGTGATG 148648





```

146847 TAGCTTGAGTCAGACAGACAACTGTCGACACTGGAAGAG 146798
408 IILCYASPRYPYRPRARYGAGIILIEALALYSTPHISALILEUCYSAS 424
|||||
146797 ATCTGTGACTACCCAGACAGACAATTGCCAAGTGCACCTGCTGTGA 146748
424 PG1Y 425
|||||
146747 TGGT 146744

seq_name: gb_htg:AC068125

seq_documentation_block:
LOCUS AC068125 170218 bp DNA HTG 24-AUG-2000
DEFINITION Homo sapiens chromosome 18 clone RP11-403A5, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
ACCESSION AC068125.5 GI:9887816
VERSION AC068125.5 GI:9887816
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 170218)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170218)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108 USA
On Aug 24, 2000 this sequence version replaced gi:9739343.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0403A05
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: Plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 159019 bases at least Q40
Consensus quality: 162717 bases at least Q30
Consensus quality: 164769 bases at least Q20
Insert size: 187000; agarose-fp
Insert size: 168718; sum-of-contigs
Quality coverage: 4.51 in Q20 bases; agarose-fp.
Quality coverage: 4.52 in Q20 bases; sum-of-contigs
----- NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 8710: contig of 8710 bp in length
* 1 8711: gap of unknown length
* 8811 18252: contig of 9442 bp in length
* 18253 18352: gap of unknown length
* 18353 27041: contig of 8689 bp in length
* 27042 27141: gap of unknown length
* 27142 33815: contig of 6774 bp in length
* 33815 34015: gap of unknown length
* 34015 43725: contig of 9710 bp in length
* 43725 43726: gap of unknown length

```

```

* 43826 56004: contig of 12179 bp in length
* 56005 56104: gap of unknown length
* 56105 70333: contig of 14229 bp in length
* 70334 70433: gap of unknown length
* 70434 83083: contig of 12650 bp in length
* 83084 83183: gap of unknown length
* 83184 99044: contig of 15861 bp in length
* 99045 99144: gap of unknown length
* 99145 135541: contig of 36497 bp in length
* 135542 135741: gap of unknown length
* 135742 137570: contig of 1829 bp in length
* 137571 137670: gap of unknown length
* 137671 141527: contig of 3857 bp in length
* 141528 141627: gap of unknown length
* 141628 147793: contig of 6166 bp in length
* 147794 147893: gap of unknown length
* 147894 155899: contig of 8006 bp in length
* 155900 155999: gap of unknown length
* 156000 162402: contig of 6403 bp in length
* 162403 162502: gap of unknown length
* 162503 170218: contig of 7716 bp in length.

FEATURES
Source
1..170218
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/clone="RP11-403A5"
1..8710
/clone="assembly_name:Contig10"
vector_side="left"
8811..18252
/clone="assembly_name:Contig11"
18353..27041
/clone="assembly_name:Contig12"
27142..33915
/clone="assembly_name:Contig13"
34016..43725
/clone="assembly_name:Contig14"
43826..56004
/clone="assembly_name:Contig15"
56105..70333
/clone="assembly_name:Contig16"
70434..83083
/clone="assembly_name:Contig17"
83184..99044
/clone="assembly_name:Contig18"
99145..135641
/clone="assembly_name:Contig19"
135742..137570
/clone="assembly_name:Contig19"
137671..141527
/clone="assembly_name:Contig19"
141628..147793
/clone="assembly_name:Contig19"
147894..155899
/clone="assembly_name:Contig19"
156000..162402
/clone="assembly_name:Contig19"
162503..170218
/clone="assembly_name:Contig19"

BASE COUNT 52643 a 30665 c 30050 g 55356 t 1504 others
ORIGIN
alignment_scores:
Quality: 1595.00 Length: 1352
Ratio: 3.853 Gaps: 4
Percent Similarity: 30.621 Percent Identity: 30.399
alignment_block:
US-09-680-121-2 x AC068125/rev ..

```

Align seg 1/1 to reverse of: AC068125 from: 1 to: 170218

```
12 AspgIuileProthrValIglIyIlePheSerAlaPheGlyLeuValPh 28
|||||
75085 GATGAAATCCACAGTGGGATCTTCAGTCATTTGGCCTGCTCTT 75036
28 eThValSerLeuphealArpIleCysCysGlnArGlySerSerLys 45
|||||
75035 CACAGTCTCTCTTTCATGATGATCTCTCTCAGAGAAATCATCAAGT 74986
45 ePAsnLysThrProProTyrLysPheValHisValLeuLysGlyValAsp 61
|||||
74985 CTACACAGACTCTCCATACAGATTGTGTCATGCTTAAGGAGTTGAT 74936
62 IleTyrProGluAsnLeuAsnSerLysLysLysPheGlyAlaAspAspLys 78
|||||
74935 ATTTACCTGAAAACCTTAATACGAAAAAGAACTTGGAGCAGATGATTA 74886
78 sAsnGlyValLysAsnLysProAlaValProLysAsnSerLeuHisLeu 95
|||||
74885 AATGAGATTAAAGATTAAGCCAGCTGTGCCAAGAAATTCATTGCACTGG 74836
95 sPLeuGluLysArGAspLeuAsnGlyAsnPheProLysThrAsnLeuLys 111
|||||
74835 ATCTTGAAAAGAGATCTCAATGCAATTTCCAAAACCAACTCAAA 74786
112 ProGlySerProSerAspLeuGluAsnAlaThrProLysLeuPheLeuGlu 128
|||||
74785 CCTGGCACTCTTCGTGATCTGGAGAAATGCACCCGCAACCTCTTTTGA 74736
128 uGlyLysGluSerValSerProGluSerLeuLysSerSerThrSerL 145
|||||
74735 AGGGAAAAAGAGTCACTTCCCTGAGATTAAAGTCCACACTTCCC 74686
145 eUrThrSerGluGluLysGlnGluLysLeuGlyThrLeuPhePheSerLeu 161
|||||
74685 TTACTTCAGAGAGAAACAAAGAGAGCTGGAACTCTCTCTCTCTTA 74636
162 GluTyrAsnPheGluArGlyLysAlaPheValIleLysGluAlaAr 178
|||||
74635 GAATTCACACTTCGAGAGAAAAGCTTTGTGTCATATCAAGAAAGCCCG 74586
178 gGlyLeuProAlaMetAspGluGlnSerMetThrSerAspProTyrIleL 195
|||||
74585 TGC.TTGCACACCATGATGAGCAGTCGATGACCTGTGACCATATATCA 74537
195 ySerMetThrIleLeuProGluLysLysHisLysValLysThrArGlyVal 211
|||||
74535 AATATACATCTCTCCAGAGAAAGATTAAGATCAAGAACTAGAGAGTCTG 74487
212 ArGlyThrIleAspProAlaPheAspGluThrPheThrPheTyrGlyI 228
|||||
74486 AGAATAAACCTTGATCCAGCTTTTGATGAGACCTTACATCTTATAGGGAT 74437
228 eProTyrThrGlnIleGlnGluLeuAlaLeuHisPheThrIleLeuSerP 245
|||||
74436 ACCCTACACCCCAATTCAGAAATTTGGCCTTGCACTTCACAAATTTGAGATT 74387
245 heAspArgPheSerArgAspAspIleIleGlyGluValLeuIleProLeu 261
|||||
74386 TTGACAGTCTTTCAGAGATGATATCATTTGGGAAAGTTCTTAATCTCTC 74337
262 SerGlyIleGluLeuSerGluGlyLysMetLeuMetAsnArGluIleI 278
|||||
74336 TCGGGAATTTGAATATCTGAGAGAAAAATGTTAATGATATAGAGATCAT 74287
278 eLysArgAsnValArg..... 283
|||||
74286 CAAGGAAATGTTAGGCTATTTATTTCAGTGCCTTAAGATTTATATAA 74237
283 ..... 283
74236 CAATCTTTTATGTAGCATTTGGTGTCTAAGAAATTTTATTAATATGAT 74187
```

```
283 ..... 283
74186 AAGCTTGATATATCATCTTACATGTTTAAATATATATATATATATC 74137
283 ..... 283
74136 TAATGAGATCAAGGAAATAGAACAGATGAATAAATTAATCATATAC 74087
283 ..... 283
74086 TGAATATCTCAAAAGTGTTTAATGATGTAATAATAAGCATTCCTTAT 74037
283 ..... 283
74036 GGATATTTAGATATCACCTATATATAATCTTTGTATCTATATCAATA 73987
283 ..... 283
73986 ATATTTTGACAGACTGCTACTTAAATAATACAGTGAAATATTTCTTAT 73937
283 ..... 283
73936 CATTAAGCAAGTTAACAAAGAGATTTCAAGTCAACATTAATGAATTTAA 73887
283 ..... 283
73886 ATTGAATTTAAATTAATTAATCAATTTCTCTGAAGATTTGGAATTTGAT 73837
283 ..... 283
73836 TGACTTTCATGGAATTTGCTTTCAAGAGAGAGCATCAGTCATTTTCA 73787
283 ..... 283
73786 ATTATAAATGCTCATCTATAATTTGTATATAGTAAGAAATGATCCCGTAT 73737
283 ..... 283
73736 TTTTGTATGTACACTACTCTGTAATCTGTGTTATTAGAAAATAGTAA 73687
283 ..... 283
73686 GACATCAATGAACAAATTTAAACCCATTTTCACTTAATTTGAACCTC 73637
283 ..... 283
73636 GAATTCACAGCTGATATATCAGAAATACCTCTCTCTATTAATACAAATGAT 73587
283 ..... 283
73586 TATTAATAATCAATAGACATTAATTTGTGTTTTTAAACAAAGATAGTGT 73537
283 ..... 283
73536 TTCAAGAGCTGAAAACCTGAGAGCTCATTTTGACTGTCTCTCTCTATTTG 73487
283 ..... 283
73486 AAATTAATAGCGTCAAGGTTCAGATACATTCACACAGCTTGTAGACA 73437
283 ..... 283
73436 CCGGCTGGTTAAGTATACAGTGGGAATTAAGATCTTTTCTACTCTT 73387
283 ..... 283
73386 ATTCTCTTAGAGTGCCTACTACTATACAGCTGCCATATATAGTTTGTGA 73337
283 ..... 283
73336 TAAATATATTCAGAGTTTATGTTTGTGTGACGCCATTCACCTCAGTTGCAA 73287
```

283 ..... 283  
72286 ACAACACCAAGTTAAATATGTACTTTCAGTCAAACTTTTACTTGG 72237  
283 ..... 283  
72326 AATTATGTATATCTTATCTATTTGTCTTAAAGTACAAATGAATCTTAG 72187  
283 ..... 283  
72186 TAGCAGCAAGAAGACGCTATGACACGTGCCAAAAACATGCTCCCA 72137  
283 ..... 283  
72136 ATGCTTAGATCAATTCATTCCTTGGCTTAAATGAAACTTGTCTGTGG 72087  
283 ..... 283  
72086 AAGAGGAGAGAAATTTTAATCCATTATCCCTGGGCCAGAAATGTAG 72037  
283 ..... 283  
72036 TTGGGATCAGAAATATTCACCTGGGTCTGAGTTCAAAATTCATATGTC 72987  
283 ..... 283  
72986 GGTGCTAAGAAATTTGAATCAATCAACTATCTAATTTTTCATGA 72937  
283 ..... 283  
72936 GTATGTAGTCAATTTGACCACACATCTCATGCAGAAATTAGATAA 72887  
283 ..... 283  
72886 GCTTTTACTATTAACTTTGAAAAGTAGAGATGTTATTAAAAACA 72837  
283 ..... 283  
72836 AAGCCAAAATATTCCTGTGTAGATATATTAATATCCTTTAACA 72787  
283 ..... 283  
72786 TTTCTAGTAATGGGTGTGGCAATTTATGCCAATTTTAACTTCGTGA 72737  
283 ..... 283  
72736 GTACCAACAGACACAGAGTGGCAGGAGTGTTCACAGCAGCAT 72687  
283 ..... 283  
72686 GGTACTGCTCAAAAGCTTTTACTCTTCAATTCCTTAGAATATTAAA 72637  
283 ..... 283  
72636 ATAAAAATAGTAGATTATATTAAATGACGATGTTATTAAACTTTAA 72587  
283 ..... 283  
72586 TGAAGTACACACTTCAAGTTAAAAAGATGAAATACTATATACACTT 72537  
284 ..... 284  
72536 TCTCTTTCGACAAAGTCTTCAGGACGGGTGAGTTACTGATCTCTCT 72487  
284 ..... 284  
72486 GCTATGAGTCCACCAACACCTTAACCTGTGTCTTAAAGCTCGA 72437  
284 ..... 284  
72436 CATCTGCTTAATCTGATGTGCGGACCTTCAGGTAAAGATCTTCTGA 72387  
283 ..... 283

72386 AATATTCAGATTCACGTATGATGAATGTATTGACTGCAAATGAGCC 72337  
323 ..... 323  
72336 CACCTACTTACTCTTCTATCTTTCTCTTAATACCTTAATGATCTGCTA 72287  
323 ..... 323  
72286 TAAAGTTAGCTTTTACTTTTATGACTAGCTTTACAAAAAATGAAACAA 72237  
323 ..... 323  
72236 AGCCAAATTTATGTGTATAGCAGCAATGAGAAACAAATAGCTACAGT 72187  
323 ..... 323  
72186 TTAAATTTACCTGATGTTCAAACTCCTAACCTTCCTCAGTACCAATA 72137  
323 ..... 323  
72136 TGATGGGTATGTTCTTAACTGTACAACTGACATAATGCTACTCAT 72087  
323 ..... 323  
72086 CCGAGGGGCTTGTGAGAACTAGAAATACATCTTCTCAGATATCATC 72037  
323 ..... 323  
72036 TAAATATATTTATTAATTAATGATGGAGTTGAAAAACATGCTTAAA 71987  
323 ..... 323  
71986 AGATTAGATTGTAGCTAAATTAATCAATATGTCACACAGAAAAAG 71937  
323 ..... 323  
71936 GCAAAAGTAAATGTGATTAAGTAGCATGAACCTTAAATCTACATCTAAC 71887  
323 ..... 323  
71886 AATGACACAGAGAGCTCAACACTAGGAAACAGAAACAAATCCAGCCTG 71837  
323 ..... 323  
71836 TGACTAAAGAAATTCATTAAGTTAAAGCTAAAGAAATTTGCCATGCT 71787  
323 ..... 323  
71786 GGAATCTAATTCCTCATTCGATGSAANAACACAGCCAAAGTCAGGGAAT 71737  
323 ..... 323  
71736 AGACTCACCCAAAGTCATACAGTCAAGTAAAGCTACATGTTGCTAATAT 71687  
323 ..... 323  
71686 ATTTTTCCTTCTATCTATCTCTGAGTAAGTAATAGTAATCAGAA 71637  
323 ..... 323  
71636 TGGTCCATAGAACATTCATTTCTTCTTCTTGTGTACTTCTTCC 71587  
323 ..... 323  
71586 TATCATTAATTTTTCCTTATCTCTAATTTAGAACTTTTCTCTTAT 71537  
323 ..... 323  
71536 TTTCATATTTCTTCTTATTTTCAATTCGTGTGACAGAGAAATGCCAAG 71487  
323 ..... 323



```

170 PheValValAsnIleuysgluaIaArgIleuProAlaMetAspGlu 186
    ::::::::::::::::::::::::::::::::::::::::::::::
517 CTGTGTGTGACATCCAGAGGCTCACGGGCTCCCATGTATGATGACCA 566
186 nserMetThrSerAspProTyrIleuysMetThrIleLeuProGluLysL 203
    ::::::::::::::::::::::::::::::::::::::::::::::
567 GACCCAGGGCTCCGATCCCTACATCAAAATGACATCTCCCGACCAAGC 616
203 ystHisValLysThrArgValLeuArgLysThrIleuAspProAlaPhe 219
    ::::::::::::::::::::::::::::::::::::::::::::::
617 GGATCGACGAGGAAGACCCGAGTCTCGCAAGACGCTGGACCCGGTCTTC 666
220 AspGluThrPheThrPheTyrGlyIleProTyrThrGlnIleGlnIle 236
    ::::::::::::::::::::::::::::::::::::::::::::::
667 GATGAGACCTTACCTTCTACGGCATCCCTACAGCCAGCTGCAGGACCT 716
236 ualaLeuHisPheThrIleLeuSerPheAspArgPheSerArgAspAsp 253
    ::::::::::::::::::::::::::::::::::::::::::::::
717 GGATCGATCTCTCTGCTGCTCAGCTTGAATCGATCTCTCGGATGACG 766
253 leilegIgluValIleuIleProLeuSerGlyIleGluLeuSerGly 269
    ::::::::::::::::::::::::::::::::::::::::::::::
767 TCATCGAGAGGTCATGCTGCTGATGCTGAGTGCAGCCAGCACAGC 816
270 LysMetLeuMetAsnArgGluIleIleLysArgAsnValArgLysSer 286
    ::::::::::::::::::::::::::::::::::::::::::::::
817 AAGGTTACGTGACCCAGGACATCATCAAGAGGACATTCAGAAAGTGCAT 866
286 rglArgGlyGluLeuLeuIleSerLeuCysTyrGlnSerThrThrAsn 303
    ::::::::::::::::::::::::::::::::::::::::::::::
867 TAGAGAGGGGAGAACCTCCAGTATCTGTATACACCCGCTGCACAGA 916
303 hLeuThrValValValLeuLysAlaArgHisLeuProLysSerAspVal 319
    ::::::::::::::::::::::::::::::::::::::::::::::
917 GAATGACAGGAGTGCTCTCAAGCCGACACTTGCAGAAAGATGATATC 966
320 SerGlyLeuSer... AspProTyrValLysValAsnLeuTyrHisAla 335
    ::::::::::::::::::::::::::::::::::::::::::::::
967 ACCGGTCTCTCAGGTATCTCTATGTCAAGTGAACGCTCTACACGGC 1016
335 sLysArgIleSerLysLysLysThrHisValLysLysCysThrProAsn 352
    ::::::::::::::::::::::::::::::::::::::::::::::
1017 AAAAGCATTCGCCAAGAAACCACCATGTAAAGATGCACCTTTGAACC 1066
352 lavalPheAsnGluLeuPheValPheAspIleProCysGluGlyLeu 368
    ::::::::::::::::::::::::::::::::::::::::::::::
1067 CCGCTTCATGAGTCTCATATGATGATCCCGACCGACCTTCTGCTCT 1116
369 AspIleSerValGluPheLeuValLeuAspSerGluArgGlySerArg 385
    ::::::::::::::::::::::::::::::::::::::::::::::
1117 GATATGACATCCAGTCTCTTGTATGACTGCTGATCGCAGCTACTAAGA 1166
385 ngluValIleGlyGluLeuValLeuGlyAlaAlaGluGlyThr...G 401
    ::::::::::::::::::::::::::::::::::::::::::::::
1167 TGAGGTGTGAGGAGGCTGATCTGTGGGCAACAGTGTACACACCAAGT 1216
401 lylGluGluHisThrPylsGluIleCysAspTyrProArgArgGlnIle 417
    ::::::::::::::::::::::::::::::::::::::::::::::
1217 GTGGGGAACCTGTGAGAGGTCTGCGAAGATCCCGCAAGCCCATAGCC 1266
418 LysTrpHisValLeuCysAsp 424
    ::::::::::::::::::::::::::::::::::::::::::::::
1267 AAGTGGCAGCTGTGAGCAG 1287

```

seq\_name: gb\_pr:BC004291

seq\_documentation\_block: 2950 bp mRNA PRI 12-JUL-2001  
 LOCUS BC004291  
 DEFINITION Homo sapiens, similar to synaptoctagmin 11, clone MGC:10881  
 IMAGE:3621175, mRNA, complete cds.  
 ACCESSION BC004291  
 VERSION BC004291.1 GI:13279139  
 KEYWORDS MGC.

```

SOURCE
ORGANISM human.
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
JOURNAL 1 (bases 1 to 2950)
DIRECT SUBMISSION Strausberg, R.
Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lissa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stolt,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 13 Row: 3 Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
FEATURES
source
Location/Qualifiers
1..2950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:10881 IMAGE:3621175"
/tissue_type="Lymph, Burkitt lymphoma"
/clone_lib="NIH_MGC_8"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
184..1479
/codon_start=1
/product="Similar to synaptoctagmin 11"
/protein_id="AAH04291.1"
/db_xref="GI:13279140"
/translation="MAETINIRSPVSPVAVGLIGASVLYVYCVSYVWSCCHQAA
EKKHKRPYKFTIHLKGISIYPTETLSNKKIKYVRDXKQPRGREGRRNLVDAAEAG
LUSRDQPRGSSGCTIDLPTRKMDGEELRSPTLSLJFGESRTISPSSEEDVWLAGS
LTFSEVDYNPVKALVYIOEANGLPVMDQTOGSDYIKMTILPDKRHVKRVLKRT
LDPVEDFTFPYGIYSQLODLVHLFVLSFPRFSDVDYIGEVMPDLACVDSTGKVO
LNRDIKRNIOKICISRGELVSLSYOPVQRMVYVYLKARHLPRKMDITGSGMPYKV
NVYGRKIRAKKTHYKCTINDIENESFTYDIPDPLDISIEFLVIDEDRTKNVEV
VGRILGASVYASGAEMHREVCESPRKVAAMHSLSER"
BASE COUNT 790 a 708 c 704 g 748 t
ORIGIN
alignment_scores:
Quality: 1162.00 Length: 437
Ratio: 3.348 Gaps: 7
Percent Similarity: 79.405 Percent Identity: 52.403
alignment_block:
US-09-680-121-2 x BC004291 ..
Align seg 1/1 to: BC004291 from: 1 to: 2950
1 MetaIaprollethrThrSerArgGluLupheAspGluIleProThrVa 17
||||| ||||||| ||| ::||| |||||

```

```

184  AAGCTGAGATCCACCATATCCAGCTAGCTTTGATGCTACCGGTGGT 233
17  lvalyllepheser]a]phesgl]leuval]p]he]r]va]l]Ser.....L 32
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234  GGCGGCTCATCGGGGCTCTGTGCTGTGGTGTCTGTCTGGTGGACCG 283
32  eup]he]a]l]a]t]r]i]le]c]y]s]c]l]n]a]r]g]l]y]s]e]r]l]y]s]e]r]a]n]l]y]s]t]h]r 48
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284  TCTTTGCTGGTCACTGCTGCACGACGACGACGACGACGACGACGAC 333
49  p]ro]p]o]t]r]y]l]y]s]p]h]e]a]l]h]i]s]v]a]l]l]e]u]l]y]s]g]l]y]v]a]l]a]s]p]i]l]e]t]r]p]o]g]l 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
334  CCACCATCAAGATTATTACATGCTCAAGCATCAAGCATCAAGCATCA 383
65  u]a]n]l]e]u]a]n]s]e]r]l]y]s]l]y]s]p]h]e].....G]l]y]a]l]a]s]p]a 77
1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384  GACCTCAAGCAAAAGAAATATCATCAAGATGCGAGAGACAAAGATG 433
77  s]p]l]y]a]s]n]g]l]v]a]l]l]y]a]s]n]l]y]s]p]r]o]l]a]v]a]l]p]r]o]l]y]s]a]n]s]e]r]l]e]u]h]i]s 93
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
434  GTCTGGGAGGAGAGTGCAGCTAGCAAGCTGTGGTGGACGACGACGAG 483
94  l]e]u]a]s]p]l]e]u]l]y]s]a]r]g]a]s]p]l]e]u]a]n]g]l]y]a]s]n]p]h]e]p]r]o]l]y]s]t]h]r]a]n]l]e 110
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
484  GCTGGCTGTACAGCGAGAC.....AAAGATCCGAG 515
110  u]l]y]s]p]r]o]l]y]s]e]r]p]r]o]s]e]r]a]s]p]l]e]u]g]l]u]a]s]n]a]l]a]t]h]r]p]r]o]l]y]s]l]e]u..... 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
516  GGGGCTGAGCTGTGATCTGTATAGACATATACCATCAAAATGAGACT 565
126  .....P]h]e]u]e]n]g]l]y]l]u]l]y]s]g]l]u]s]e]r]v]a]l]s]e]r]p]r]o]l]u]s]e]r]l]e]u 139
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
566  AAGGGAAGAACTAAGAGCCCTATTACAAGCTGACCCCTGGGAGAGC 615
140  l]y]s]e]r]S]e]r]t]h]r]S]e]r]l]e]u]t]h]r]S]e]r]g]l]u]l]y]s]g]l]u]l]y]s]e]n]g]l]y]t]h 156
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
616  AAA.....ACCACCTCTCATCATCTCCAGAGAGAGATGATCATCTGGAGTC 662
156  r]l]e]u]p]h]e]S]e]r]l]e]u]g]l]u]t]r]y]a]s]n]p]h]e]g]l]u]a]r]g]l]y]a]l]a]p]h]e]v]a]l]a 173
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
663  CCTCACCTTCTAGTGCATATACTCCGAAAAAAGCCCTGGTGGTGA 712
173  s]n]l]l]y]s]g]l]u]a]l]a]r]g]l]y]l]e]u]p]r]o]l]a]k]e]t]a]s]p]g]l]u]n]s]e]r]t]h]r 189
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
713  CAATCAGAGAGGCCAGCGCTGCATGATGATGATGATGATGATGATGAT 762
190  S]e]r]a]s]p]p]r]o]t]r]y]l]e]t]y]s]m]e]t]h]r]l]l]e]u]p]r]o]g]l]u]l]y]s]t]h]r]a]n]l]y]s]v]a 206
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
763  TCTGACCCCTCATCAAAATGACCATCTCTCGACCAAGGCGCATCGGCT 812
206  l]y]s]t]h]r]a]r]g]l]y]a]l]e]u]a]r]g]l]y]s]t]h]r]l]e]u]a]s]p]r]o]l]a]p]h]e]a]s]p]i]l]e]t]r]p] 223
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
813  GAAAGACAGAGTGTGCGGAAGACCCGAGACCTCTGTTGACGAGACT 862
223  h]e]t]h]p]h]e]t]r]y]l]l]e]p]r]o]t]r]y]t]h]l]l]e]n]g]l]u]l]e]u]a]l]e]u]h]i]s 239
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
863  TCACCTTATATGACATCCCTACAGCAGGTGAGAGATGCTGTGAC 912
240  p]h]e]t]h]r]l]l]e]u]s]e]r]p]h]e]a]s]p]a]r]g]p]h]e]s]e]r]a]r]g]a]s]p]a]s]p]i]l]e]t]r]p]i]l]e]g]l]y] 256
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
913  TTCCTGTCCTACCTTTCACCTCTCTCGGAGATGATGATGAGCGA 962
256  u]v]a]l]l]e]u]l]l]e]p]r]o]l]e]u]s]e]r]g]l]y]l]l]e]g]l]u]l]e]u]s]e]r]g]l]y]s]m]e]t]l]e]u]m 273
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
963  GGTCTATGTGCTGCTGAGGCTGAGGCCACGACAGGCAAGGATGACAC 1012
273  e]t]a]s]a]r]g]l]u]l]l]e]l]y]a]r]g]a]s]n]v]a]l]a]r]g]l]y]s]e]r]l]y]s]e]r]g]l]y]a]r]g]l]y 289
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1013  TGACACGAGACATCATCAAAAGAAATATCCAGAGTGCATGACGAGAG 1062
290  g]l]u]l]e]u]l]l]e]s]e]r]l]e]u]c]y]s]t]r]y]g]l]n]s]e]r]t]h]r]a]n]l]y]s]t]h]r]l]e]u]t]h]r]v]a 306
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1063  GAGCTCCAGGTGTCTGTGATATACAGCTGTGGGACGACGAGAAATGAC 1112

```

```

306  lvalleu]l]y]s]a]l]a]r]g]l]h]i]s]l]e]u]p]r]o]l]y]s]e]r]a]s]p]a]s]p]i]l]e]t]r]p]i]l]e]u]s 323
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1113  GTGTGCTCTCAAAAGCCAGACCTTCCGACAGATGATGATATACCGCTCTCT 1162
323  e]r]...A]s]p]p]r]o]t]r]y]l]y]s]v]a]l]a]s]n]l]e]u]t]h]r]a]n]l]y]s]a]r]g]l]y 338
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1163  CAGGTATCTCTTATGTCAGAGTGAAGCAGTCTACTACGCGACCAAGCCGAT 1212
339  S]e]r]l]y]s]t]h]r]h]i]s]v]a]l]l]y]s]c]y]s]t]h]r]p]r]o]a]n]a]l]a]l]p]h]e]a]s 355
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1213  GCCAAGAAAGAAACCCTGTGAAGAAAGTGCACCTTGAACCCCATCTTCAA 1262
355  n]g]l]u]l]e]u]p]h]e]v]a]l]p]h]e]a]s]p]i]l]e]p]r]o]c]y]s]g]l]u]l]y]l]e]u]g]l]u]s]p]l]e]s]e]r]v 372
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1263  TGAATCTTTCATCTACGACATCCCATGACCTGCTGCTGCTGATGACGA 1312
372  a]l]g]l]u]p]h]e]u]v]a]l]l]e]u]a]s]p]e]r]g]l]u]a]r]g]l]y]s]e]r]a]r]g]a]s]n]g]l]u]l]l]e 388
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1313  TCGAGTCTCTGCTTATGACATTCGATCGACGACGACGACGACGACGAGT 1362
389  g]l]y]g]l]u]l]e]u]v]a]l]l]e]u]g]l]y]a]l]a]...A]l]a]l]a]g]l]u]l]y]h]r]g]l]y]g]l]u]h]i 404
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1363  GGGAGCTGATCTCTGGGGGACACAGTGTACAGCCAGTGTGTGTGACGA 1412
404  s]t]r]p]l]y]g]l]u]l]l]e]c]y]a]s]p]t]r]p]r]o]a]r]g]l]u]l]l]e]a]l]y]s]t]r]p]h]i]s]y 421
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1413  CTGAGAGAGAGTCTGCGAGAGCCCGCCGACAGCTGTGGCAAGTGGACA 1462
421  a]l]l]e]u]c]y]a]s]p 424
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1463  GTCTGAGCGAG 1473

seq_name: gb_ro:AF375465
LOCUS AF375465 1293 bp mRNA ROD 27-MAY-2001
DEFINITION Rattus norvegicus synaptotagmin 11 mRNA, complete cds.
ACCESSION AF375465
VERSION AF375465.1 GI:14210271
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1293)
von Poser, C., Ichtenko, K., Shao, X., Rizo, J. and Sudhof, T.C.
The evolutionary pressure to inactivate A subtypes of
synaptotagmins with an amino acid substitution that abolishes Ca2+
binding. J Biol. Chem. 272 (22), 14314-14319 (1997)
MEDLINE 97306241
PUBMED 9162066
2 (bases 1 to 1293)
Sudhof, T.C.
Direct Submission
Submitted (02-MAY-2001) Center for Basic Neuroscience, Department
of Molecular Genetics, and Howard Hughes Medical Institute,
University of Texas Southwestern Medical Center, 6000 Harry Hines
Bldv., Dallas, TX 75390-9111, USA
LOCATION/Organism
1. 1293
/db_xref="taxon:10116"
/organism="Rattus norvegicus"
1. 1293
/note="putative membrane trafficking protein"
/codon_start=1
/product="synaptotagmin 11"
/protein_id="AAK56960.1"
/db_xref="GI:14210272"
/translation="MAEITNIRSPDVSFAVAGLIGASVLYVCSVTIVFVWTCDDQA

```

334 a 334 c 340 g 275 t  
 GRLLGASHVTTSGAEHMRVECSRRPRVAAMHSLSEY"

```
alignment_scores:
  Quality: 1160.50
  Ratio: 3.413
  Percent Similarity: 76.749
  Length: 443
  Gaps: 7
  Percent Identity: 51.693
```

```
alignment_block:
US-09-680-121-2 x AF375465
```

Align seg 1/1 to: AF375465 from: 1 to: 1293

1 MetAlaProIleThrThSerArgLysLysLysAspGluLeuProThrVal 17  
117 ValGlyLeuPheSerAlaPheGlyLeuValPheThrValSer.....L 32  
51 GGCTGACCTCATCGGCGCCTCGTGTGCTGTGTGTGTCTGTGTACTG 100  
1 ATGCTGAGATCAAAATATACGAGCCCAAGCTTTATATGTCTACCAAGAGC 50  
171 ValGlyLeuPheSerAlaPheGlyLeuValPheThrValSer.....L 32  
51 GGCTGACCTCATCGGCGCCTCGTGTGCTGTGTGTGTCTGTGTACTG 100  
32 eufPheAlaTrpIleCysGlnArgLysSerSerLysSerAsnLysThr 48  
101 TCTTTGTGTGGACATGCTGCCACACGAGCGGAGGAAGAAGACACAGAGCC 150  
49 ProProTyrLysPheValHisValLeuLysGlyValAlaSplIleYrProG 65  
151 CCACCTACAAAGTTATTCATATGCTGTGAAGAGCATTCAGCATCCAGCA 200  
65 uAsnLeuAsnSerLysLysLysPheGlyValAspAspLysAsnGluValL 82  
201 GACCTTCAGCACAAAGAAAGAAATCATCAAGTTCCGGAGGGACAAAGATG 250  
82 yAsnLysProAlaValProLysAsnSerLeuHisLeuAspLeuGluLys 98  
251 GTTCTCATATGGGGAAGATGGAGACAGAGGAACCTGTATGATAATGCGAGTCT 300  
99 .....ArgAspLeuAsnGlyAsnPro..... 106  
301 GGCCTCTGACCATGACATGAGGATCCCAAGAGGGCTACGCTGCATCTTG 350  
107 .....LysThrAsnLeuLysProGlySerProSerA 117  
351 TATAGACCAAGTTACCCATCCAAAGAGACTATGGGGAAGAACTGAGAGGCC 400  
117 sPLeuAlaAsnAlaThrProLysLeuPheLeuGlnGlyLysGluSer 133  
401 CCATGACCAAGCTGACCC.....GGTGAAGCAAGCC 435  
134 ValSerProGluSerLeuLysSerSerThrSerLeuThrSerGluGly 150  
436 ACCTGCCCA.....TCATCGCAGCAAGCA 458  
150 sGlnGluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluA 167  
459 AGATGTGATGTAGAGGTCCCTTACCTTCCGTGTGACTACAACTTTCCCA 508  
167 rGlyAlaPheValValAsnLysGluAlaArgGlyLeuProAlaMet 183  
509 AGAAAGCTGTGGTGGATCAATCCAAAGGCGCATGGGCTGCCAGTATG 558  
184 AspGluGlnSerMetThrSerAspProTyrIleLysMetThrIleLeuP 200  
559 GATGGCCAAACCCAGGCGCTCATCTCCATCAATCAATGACATTCCTCC 608

EKRRKTPPYEYFIHMLKGISIYPERLSNKKKILIVRDKDQSHSESGRGLLVNAGSL  
 LSHRDQSGSPAGCIDOLPIKRDYGEILRSPMTSLTPESKPTSSPSPEEDVMGSL  
 TFSVDYNEPRKALVYTIQEAHAGLPWMDQOTQSDPIKMTILPDKHRKTYLVKRTL  
 DPVDFEFTFGLYSOLOLVLPFLVDFRSDVDVIGEVWVPLAGVDPSTGVOL  
 TRDIKFNIRKQICISRGELQVLSYOPVAKRMIVLVKARHILPRMDITGLSGMPYKVN  
 VYVRKIKIAKTKTKLPIENESIYDITDPLDITSEIFIVIDPDRTTKNEVV  
 GRLLIGASVTTSGAEHMRVESPSPRPVAKHMSLEY"

COUNT 334 a 344 c 340 g 275 t

Comment\_scores:  
 Quality: 1160.50 Length: 443  
 Ratio: 3.413 Gaps: 7  
 Percent Similarity: 76.749 Percent Identity: 51.693

Comment\_block:  
 09-680-121-2 x AF375465 ..

gln seg 1/1 to: AF375465 from: 1 to: 1293

1 MetalProIIeThrThySerArgGluIupheAspGluIleProThyVal 17  
 ||||| |||||::: ||| ::||| |||||::: |||||  
 1 ATGGCTGAGATCTCAAAATATACGCCGCCAGCTTTGATGTGTCTCACAGTACG 50  
 17 ValGluIlePheSerAlaPheGlyLeuValPheThyValSer .....L 32  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 51 GCGTGGCGCTCATCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 100  
 32 eubhealTrpIleCysGlnArgIleGlySerSerIleCysSerAsnIleThr 48  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 101 TCTTTGTGTGACATGCTGCCACGACGAGCGGAGAGACAGACAAAGACC 150  
 49 ProProGlyThrPheValHisValIleuSerGlyValAspIleIleuProGly 65  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 151 CCACGCTACCAAGTTTATTCATATGCTGAAGGCAATTCATTCACCTACCCAGA 200  
 65 uasnLeuAsnSerIleCysIlePheGlyAlaAspIleCysAsnGluValL 82  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 201 GACCCGTCAGAACAAAGAAATATCTCAAAAGTTGGAGGAGCAAAAGATG 250  
 82 yasnIleProAlaValProIleAsnSerLeuHisIleuAspLeuGluIleCys 98  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
 251 GTTCTCATAGGGAAGTAGGAGGAGAACCTGTAGTAATATCGGAGTCT 300  
 99 .....ArgAspLeuAsnGluAsnIlePro..... 106  
 301 GGCCTCTCGAGCCATGACAGAGGATCCCAAGAGGCGCTAGCGCTGCATCTTG 350  
 107 .....LysThrAsnIleuSerProGlySerProSerA 117  
 ||| ::||| |||||  
 351 TATAGACCAAGTTACCCATCAAAAGAGACTATGGGAGAGAACAGACGAGACC 400  
 117 splenGluAsnAlaThrProIleCysLeuPheLeuGluGlyIleCysGluSer 133  
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

seq\_name: gb\_pr:AK027540

seq\_documentation\_block:

LOCUS	AK027540	2315 bp	mRNA	PRI	15-MAY-2001
DEFINITION	Homo sapiens cDNA FLJ14634 fis, clone NT2RP2001081, moderately similar to SYNAPTOTAGMIN IV.				

AK027540.1

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE  
1 (sites)

Nishikawa, T.





[illegible]

1  
:  
:  
:





alignment\_scores:  
 Quality: 2201.00 Length: 425  
 Ratio: 5.179 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-09-680-121-2 x AAV54208 ..

Align seg 1/1 to: AAV54208 from: 1 to: 3891

```

1 MetAlaProTleThrThrSerArgGluIupheAspGluIleProThrVal 17
  |||||
151 ATGGCTCGGATCCACCGACCGGAAAGATTGATGAATATCCACAGT 200
  |||||
17 ValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPheA 34
  |||||
201 GGTGGGATCTTCAGTGCATTTGGCTGGCTTCACAGTCTCTCTTTG 250
  |||||
34 lATrPIleCysCysGlnArgLysSerSerLysSerAsnLysThrProPro 50
  |||||
251 CATGATCTGCTGCAGAGAAATCATCCAAAGCTAACAAGACTCTTCCA 300
  |||||
51 TyrLysPheValHisValLeuLysGlyValAspIleTyrProGluAsnLe 67
  |||||
301 TACAGATTGTGCATGTGCTTAAGGGAGTTGATATTACCTGGAAMACCT 350
  |||||
67 uAsnSerLysLysLysPheGlyAlaAspAspLysAsnGluValLysAsnL 84
  |||||
351 AAATGCAAAAAAGAGTTGGAGCAGATGATAAAATGAAGTAAGAATA 400
  |||||
84 ySProAlValProLysAsnSerLeuHisLeuAspLeuGluLysArgASP 100
  |||||
401 ACCGAGCTGTGCCAAAGAAATTCATTCATCTGCATCTTGAAAGAGAGAT 450
  |||||
101 LeuAsnGluAsnPheProLysThrAsnLeuLysProGlySerProSerAs 117
  |||||
451 CTCATGGCAATTTTCCCAAAACCACTCAAACTGGCAGTCTTTCGA 500
  |||||
117 PleuGluAsnAlaThrProLysLeuPheLeuGluGluLysGluSerV 134
  |||||
501 TCTGGAGAAATGCCAACCCGAAAGCTCTTTTAGAAGGGGAAAAAGAGTCA 550
  |||||
134 aLSeRProGluSerLeuLysSerSerThrSerLeuThrSerGluLys 150
  |||||
551 TTTCCCTGAGAGTTTAAGTCCAGCACTTCCCTTACTTCCAGAGAGAA 600
  |||||
151 GlnGluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluAr 167
  |||||
601 CAAGAGAACCTGGGAACTCTCTCTCTCTAGAAATACAACTTCGAGAG 650
  |||||
167 gLysAlaPheValValAsnIleLysGluAlaArgGlyLeuProAlaMetA 184
  |||||
651 AAAAGCATTTGGTCAATATCAAGAAAGCCGTGGCTTGCCAGCCATCG 700
  |||||
184 sPGLuGlnSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200
  |||||
701 ATGAGCAGTCGATGACCTCTGACCATATATCAAAATGAGAGATCTCCA 750
  |||||
201 GlnLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217
  |||||
751 GAGAGAACCATTAAGTGAACCTAGAGTCTGAGAAAAAAGCTTGATCC 800
  |||||
217 oAlaPheAspGluThrPheThrPheTyrGlyIleProTyrThrGlnIleG 234
  |||||
801 AGCTTTTGATGAGACCTTACATTCATGGATACCTTACACCCAAATCC 850
  |||||
234 InGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250
  |||||
851 AACAAATTGGCCTTGCACTTCACAAATTTGAGTTTGACAGGTTTCAAGA 900
  |||||
251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeuSe 267
  |||||
901 GATGATATCATTTGGGAAAGTTCTAATTCCTCTCTCGGAAATTAATATC 950
  |||||

```

```

267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284
  |||||
951 TGAAGAGAAAAATGTTAAATGAATAGAGATCATCAAGAGAAATGTAGGA 1000
  |||||
284 ySerSerGlyArgGlyGluLeuLeuIleSerLeuCysTyrGlnSerThr 300
  |||||
1001 AGCTCTCAGAGACGGGTGAGTTACTGATCTCTCTCTATCATAGTCCACC 1050
  |||||
301 ThrAsnThrLeuThrValValValLeuLysAlaArgHisLeuProLysSe 317
  |||||
1051 ACNAAACACTCACTGCTGTGTTCTTAAAGCTCGACATCTGCTTAAATC 1100
  |||||
317 rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuTyrHisA 334
  |||||
1101 TGATGTCTCGGACTTTCAGATCCCTATGTCAAAAGTGAACCTGTACATG 1150
  |||||
334 lAluLysArgIleSerLysLysThrHisValLysLysCysThrPro 350
  |||||
1151 CCAAAAGAGAAATCTCCAGAGAAAGACTCATGTGAAGAAATGCACCCCC 1200
  |||||
351 AsnAlaValPheAsnGluLeuPheValPheAspIleProCysGluGlyLe 367
  |||||
1201 AATGCAGTGTCAATGAGCTGTTGTCTTGATATATCTTGATGAGGACCT 1250
  |||||
367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgGlySera 384
  |||||
1251 TGAAGATATTAAGTGTGAATTTTGTGTTTGATTCGAAAGGGGTCC 1300
  |||||
384 rGAsnGluValIleGlyGlnLeuValLeuGlyAlaAlaGluGlyThr 400
  |||||
1301 GAATAGAGGAATCGGCGAGTTAGCTTGGGTGCAGCAGCAAGAAAGACT 1350
  |||||
401 GLyGlyGluHisTrpLysGluIleCysAspTyrProArgArgGlnIleAl 417
  |||||
1351 GGTGGAGACACTGGAAAGAGATCTGTGACTACCCCGAGAGACAAATTC 1400
  |||||
417 aLysTrpHisValLeuCysAspGly 425
  |||||
1401 CAAGTGGCAGTCTCTGTGATGTG 1425
  |||||

```

seq\_name: /sids8/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV57327

seq\_documentation\_block:  
 ID AAV57327 standard; cDNA; 3891 BP.  
 XX  
 AC AAV57327;  
 XX  
 DT 21-DEC-1998 (first entry)  
 XX  
 DE Hormone-regulated Repro-PC-1.0 gene.  
 XX  
 KW Repro-PC-1.0; prostate cancer; LNCaP; hormone-regulated gene;  
 human; synaptotagmin; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH key  
 FT CDS 151..1428  
 FT /\*tag= a  
 XX  
 PN WC9839661-A1.  
 XX  
 PD 11-SEP-1998.  
 XX  
 PE 06-MAR-1998; 98WO-US04519.  
 XX  
 PR 15-MAY-1997; 97US-0047811.  
 PR 07-MAR-1997; 97US-0041246.  
 XX  
 PA (REPR-) REPROGEN INC.  
 XX  
 PI French CK, Yamamoto KK;

XX WPI: 1998-506379/43.  
 DR P-PSDB: AAW75782.

XX Identification of hormone-regulated traits, e.g. in cancers - by  
 PT exposing grafts of biological material to different hormonal  
 XX environments in animals of different reproductive states

PS Example 1; Page 59-62; 85pp; English.

CC This nucleotide sequence represents a hormone-regulated gene from  
 CC human prostate cancer cells. In order to isolate sequences that  
 CC are over-expressed in male LNCap tumours, a male-LNCap-specific  
 CC probe was generated by 3 rounds of subtractive hybridisation with  
 CC female LNCap tumour cDNA. The probe was used to perform a primary  
 CC screen of a lambda-ZAP-male-LNCap tumour cDNA library. Positive  
 CC plaques were subjected to secondary and then tertiary screens using  
 CC male- and female-specific probes to isolate partial clone  
 CC Repro-PC-1.0. Subsequent screening of the male-LNCap tumour  
 CC library and RACE-PCR yielded a sequence containing a single 1275  
 CC open reading frame encoding 425 amino acids (see AAW75782). The  
 CC encoded protein has regions of homology to the C2 regulatory  
 CC domain of calcium-dependent isoforms of protein kinase C and to  
 CC isoforms of synaptotagmin. The gene was localised to chromosome  
 CC 18. The invention provides methods of identifying hormone-regulated  
 CC traits in a cell. The methods involve cultivating the cell as a  
 CC graft in 2 different hormonal environments and determining whether  
 CC expression of the trait differs in the 2 grafts. The methods can  
 CC be used to identify hormonally-regulated traits and hormonally-  
 CC regulated genes for use as targets for therapeutic intervention in  
 CC disease states, particularly cancers.

XX Sequence 3891 BP; 1229 A; 671 C; 788 G; 1203 T; 0 other;

# alignment\_scores:

Quality: 2201.00 Length: 425  
 Ratio: 5.179 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-09-680-121-2 x AAV57327 ..

Align seg 1/1 to: AAV57327 from: 1 to: 3891

1 MetAlaProIleThrThrSerArgGluGluPheAspGluIleProThrVal 17  
 |||||||  
 151 ATGGCTCCGATCCACCCAGCCGGGAAGAAATTGATGAATCCCCACAGT 200  
 17 lValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPheA 34  
 |||||||  
 201 GGTCGGGATCTTCAGTCATTTGGCTGGCTTTCACAGTCTCTCTTTTG 250  
 34 lATrPLeuCysGlyNArgLysSerSerLysSerAsnLysThrProPro 50  
 |||||||  
 251 CATGGATCTGCTGTCAGAGAAATCATCCAGTCTAACACAGACCTCTCA 300  
 51 TyrLysPheValHisValLeuLysGlyValAspIleTyrProGluAsnLe 67  
 |||||||  
 301 TACAAGTTTGTCATGTGCTTAAGGAGTGTATTTTAACTCTGAAACCT 350  
 67 uAsnSerLysLysLysPheGlyAlaAspAspLysAsnGluValLysAsnL 84  
 |||||||  
 351 AAATAGCAAAAGAAAGTTTGGAGCAGATGATAAATGAAGTAAAGAAATA 400  
 84 ySProAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100  
 |||||||  
 401 AGCCAGCTGTGCCAAAGAAATCATTTGATCTTGAAAGAGAGAT 450  
 101 LeuAsnGlyAsnPheProLysThrAsnLeuLysProGlySerProSerAs 117  
 |||||||  
 451 CTCATATGGCAATTTTCCCAAAACCAACCTCAACCTGGCAGTCTTCTGA 500

117 pLeuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGlySerV 134  
 |||||||  
 501 TCTGGAGAAATGCACACCCGAAGCTCTTTTAGAAGGGGAAAAAGAGTCAG 550  
 134 aLSerProGluSerLeuLysSerSerThrSerLeuThrSerGluGluLys 150  
 |||||||  
 551 TTTCCCTGGAGAGTTTAAAGTCAGCACTTCCCTTACTTACAGAAAGAAA 600  
 151 GluGluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluAr 167  
 |||||||  
 601 CAAAGAGAGCTGGGAACTCTTCTCTCTTCCCTTAGAATACACATTCGAGAG 650  
 167 gLyAlaPheValAlaAsnIleLysGlyAlaArgGlyLeuProAlaMetA 184  
 |||||||  
 651 AAAGCAATTTGTGTCATATACAGAGAAAGCCGTGGCTTGGCACCATTGG 700  
 184 sPGluInSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200  
 |||||||  
 701 ATGACAGTCGATGACCTTGACCACATATATCAAAATGACAGATCTCCCA 750  
 201 GluLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217  
 |||||||  
 751 GAGAAAGAGCATTAAGTGAAGTGAAGTGTGAGAAAAACCTTGATCC 800  
 217 oAlaPheAspGluThrPheThrPheTyrGlyIleProTyrThrGlnIleG 234  
 |||||||  
 801 AGCTTTTGATGAGACCTTTACATTTCTATGGATACCTACACCCCAATTC 850  
 234 InGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArg 250  
 |||||||  
 851 AAGAATGGCCTTGACACTTCACATTTTGAAGTTTGCAGAGTTTTCAGA 900  
 251 AspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGluLeuSe 267  
 |||||||  
 901 GATGATATCATTTGGGGAAGTTCTAATTCCTCTCGGGAATTTGAATATC 950  
 267 rGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnValArgL 284  
 |||||||  
 951 TGAAGGAAAAATGTTAATGATACAGAGATCATCAGAGAAATGTTAGGA 1000  
 284 ySSerSerGlyArgGlyGluLeuLeuIleSerLeuCysTyrGlnSerThr 300  
 |||||||  
 1001 AGTCTTCAGACGGGGGTGACTTACTGATCTCTCTGCTATCAGTCCACC 1050  
 301 ThrAsnThrLeuThrValValLeuLysAlaArgHisLeuProLysSe 317  
 |||||||  
 1051 ACAAACTCTTAACCTGTGCTTCTTAAAGCTGCACATTCGCCCTAAATC 1100  
 317 rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuTyrHisA 334  
 |||||||  
 1101 TGATGTGTCGGGACTTTCAGATCCCTATGTCAAAGTGAACCTGTATCCATG 1150  
 334 lAlLysLysArgIleSerLysLysLysThrHisValLysLysCysThrPro 350  
 |||||||  
 1151 CCAAAAGAGAGATCTCCAAAGAGAGACTCATGTGAAGAAATGCAACCCCC 1200  
 351 AsnAlaValPheAsnGluLeuPheValPheAspIleProCysGluGlyLe 367  
 |||||||  
 1201 AATGAGAGTTCAAATGAGCTGTTTGTCTTGATATCTCTTGAGGGCT 1250  
 367 uGluAspIleSerValGluPheLeuValLeuAspSerGluArgLysEra 384  
 |||||||  
 1251 TGAAGATATAGTGTGTGAATTTTGGTTTGGATTCTGAAAGGGGGGTCC 1300  
 384 rGAsnGluValIleGlyGlnLeuValLeuGlyAlaAlaIleGluGlyThr 400  
 |||||||  
 1301 GAAATGAGAGTAAATCGGCGAGTTAGTCTTGGGTGCAGCGCAGAGGAACT 1350  
 401 GLyGlyGlnHisThrLysGluIleCysAspTyrProArgArgGlnIleAl 417  
 |||||||  
 1351 GGTGAGAGCACTGGAAAGAGATCTGTGACTAACCCAGGAGACAAATTCG 1400  
 417 aLysTrpHisValLeuLysCysAspGly 425





```

496 GCGGGCTGTAGCCGAGAC.....AAAGATCCAG 527
110 uylsProglSerProSerAspLeuGluAsnAlaThrProLysLeu.... 125
528 GGGGCTTACCTTGATCTGTATAGACAAATTACCCATCAAAATGAGACT 577
126 .....PheLeuGluGlyLysGluSerValSerProGluSerLeu 139
578 ATGGGGAAGAACTAAGAGCCCTATTACAAAGCCCTGAGGAGGC 627
140 LysSerSerThrSerLeuThrSerGluGluLysGluLysGluLys 156
628 AAA...ACCACCTCTCCATCATCTCCAGAGAGATGTCATGTCAGATC 674
156 rleuPhePheSerLeuGluTyrAsnPheGluAlaGlyAlaPheValAla 173
675 CCTCACCTTCTCAGTGACATACTCCGAAAAAGCCCTGATGTGCA 724
173 snlleLysGluAlaArgGlyLeuProAlaMetAspGluInsSerMetThr 189
725 CAATCCAGAGAGCCCGGCGCTGCAGTGTATGATCAGCAGCCAGGGA 774
190 SerAspProTyrIleLysMetThrIleLeuProGluLysLysHisLysVa 206
775 TCTGACCCCTACATCAAAATGACATCTCTCTGACAAAGGCAATCGGAT 824
206 LysThrArgValLeuArgLysThrLeuAspProAlaPheAspGluThrP 223
825 GAAGACCCAGAGTGGCGGAGAGACCTGGACCTGTGTTGACGAGACT 874
223 heThrPheTyrGlyIleProTyrThrGlnIleGlnIleLeuAlaLeuHis 239
875 TCACCTTATGATGCATCCCTACAGCCAGCTGCAGACCTGTGTCGAC 924
240 PheThrIleLeuSerPheAspArgPheSerArgAspPheIleIleGlyL 256
925 TTCCTGTGCTCAGCTTGTGACCTTCTCTCGGATGATGATCATGTGCGA 974
256 uValLeuIleProLeuSerGlyIleGluLeuSerGluGlyLysMetLeuM 273
975 GGTCAATGGTGCACCTGGCAGGCGGCGCCAGCACAGCAGAGTACAC 1024
273 eTAsnArgGluIleLysArgAsnValArgLysSerSerValArgLys 289
1025 TGACCGAGGACATCATCAAAAGATATTCAGAAATGATGATCGACGAGG 1074
290 GluLeuLeuIleSerLeuCySTyrGlnSerThrThrAsnThrLeuThVa 306
1075 GAGCTCCAGAGTGTCTCTGATCATCATCACTGTGGCAGACAGATACAGT 1124
306 lValValLeuLysAlaArgHisLeuProLysSerAspValSerGlyLeuS 323
1125 GGTGTCTCTCAAGCCAGACACTTGCAGAGATGATATCACCGGTCTCT 1174
323 er...AspProTyrValLysValAsnLeuTyrHisAlaLysLysArgIle 338
1175 CAGGTAACTCTTATGTCAAGGTGACGTCTACTACGCGCAAAAGGCGATT 1224
339 SerLysLysLysThrHisValLysLysCySThrProAsnAlaValPheAs 355
1225 GCCAAGAAAGAAACCCATGTGAAAGAGTGCACCTTGAACCCCATCTTCA 1274
355 nGluLeuPheValPheAspIleProCysGluLysLeuGluAspLieserV 372
1275 TGAATCTTTCATCTACGACATCCCACTGACCTCCGCTGATATACACA 1324
372 aGluPheLeuValLeuAspSerGluArgLysSerArgAsnGluValAla 388
1325 TCGAGTTCCTGTTATGATGATGATGATGATGATGATGATGATGATG 1374
389 GlyLeuLeuValLeuGlyAla...AlaAlaGluGlyThrGlyGlyGluH 404
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

1375 GGGAGCTGATCTCGGGGCGACACAGTGTACACGACCACTGTGCTGACACA 1424
404 sTrpLysGluIleLysAspTyrProArgArgGlnIleAlaLysTrpHisV 421
1425 CTGGAGAGAGGCTCTGGAGAGCCCGGCAAGCTGTGGCCAAAGTGCGACA 1474
421 alleuCysAsp 424
1475 GTCTGAGCCGAG 1485
seq_name: /SID8/gcdata/geneseq/geneseqn/NA2001.DAT:AA160519
seq_documentation_block:
ID AA160519 standard; cDNA; 5517 BP.
XX
AC AA160519;
XX
DE 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4508.
XX
KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 13-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
XX
PT P-PSDB: AAM41363.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1: SEQ ID NO 4508; 10078bp: English.
XX
XX
The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with noctropic.
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.

```



CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX  
 S0 Sequence 5517 BP; 1480 A; 1274 C; 1294 G; 1469 T; 0 other;

# alignment\_scores:

Quality: 1154.00 Length: 437  
 Ratio: 3.335 Gaps: 7  
 Percent Similarity: 79.176 Percent Identity: 52.174

# alignment\_block:

US-09-680-121-2 x AA160519 ..

Align seg 1/1 to: AA160519 from: 1 to: 5517

```

1 MetAlaProIleThrThrSerArgGluGlnPheAspGluIleProThrVa 17
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
519 ATGGCTGAGATCACAATATCCGACCTTGGCTTGCATGCTCACCAGTGT 568
17 ValGlyIlePheSerAlaPheGlyLeuValPheThrValSer.....L 32
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
569 GGGCGGCTCATCGGGCTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 618
32 eufheAlaTrpIleCysCysGlnArgIlySerSerIlySerAsnIlyThr 48
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
619 TCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 668
49 ProProTylIlySerPheValHisValIleuLysGlyValAspIleTyrProG1 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
669 CCACCATACAAAGTTTATTCACATGCTCAAGGATCAGCATATACCCAGA 718
65 uAsnIleAsnSerIlyLysIlySrphe.....GlyAlaAsp 77
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 GACCTTGAGACACAGAAAGAAATCATCAAGTGGAGACACAAAGATG 768
77 sPluAsnGluValIlyAsnIlySrpheAlaValProLysAsnSerIleHis 93
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
769 GTCTGTGGAGAGAGAGTGGACGTAGAGAACTGTGTGTGTGTGTGTGT 818
94 LeuAspLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLe 110
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
819 GCTGGCTCTCAAGCCGAGAC.....AAAGATCCAG 850
110 uLysProGlySerProSerAspLeuGluAsnAlaThrProLysLeu... 125
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
851 GGGGCTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
126 .....PheLeuGluGlyLysGluSerValSerProGlySerLeu 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
901 ATGGGAGAAAGACTAGAGCCCTATTACAAAGCTGAGCCCTGGGAGACG 950
140 LysSerSerThrSerLeuThrSerGluGluLysGlnIlyLysLeuGlyTh 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
951 AAA...ACACATCTCCATCATCTCCAGAGAGATGTCTAGATG 997
156 rLeuPhePheSerLeuGluTyrAsnPheGluArgLysAlaPheValAla 173
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
998 CCTACCTTCTCAGTACATATGACTCCGAAAAAAGCCCTGTGTGTGA 1047
173 snIleLysGluAlaArgGlyLeuProAlaMetAspGluGlnSerMetThr 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1048 CAATCCAGAGAGCCGACGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGT 1097
190 SerAspProTylIlyLysMetThrIleLeuProGluLysLysHisLysVa 206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1098 TGTGACCCCTACATCAAAATGACATCTCTGTGACAAAGGATCGGCT 1147
206 LysThrArgValLeuArgLysThrIleuAspProAlaPheAspIlyThr 223
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1148 GAAGACCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1197
223 heThrPheTyrGlyIleProTylThrGlnIleGlnIlyLeuAlaLeuHis 239

```

```

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1198 TCACTTCTATGGATCCCTACAGCCAGCTGAGACCTGTGTCTCAC 1247
240 PheThrIleLeuSerPheAspArgPheSerArgAspIleIleGlyI 256
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1248 TTCCTTCTCTCAGCTTGTGACCGCTTCTGTGTGTGTGTGTGTGTGT 1297
256 uValIleIleProLeuSerGlyIleGluLeuSerGluGlyLysMetLeu 273
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1298 GGTATGTGTCTGACAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGT 1347
273 eAsnArgGluIleIleLysArgAsnValArgLysSerSerGlyArgGly 289
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1348 TGACCAAGGACATCATCAAAAGAAATATCCAGAGATGATGACAGAGG 1397
290 GluLeuLeuIleSerLeuCysTyrGlnSerThrThrAsnThrIleThrVa 306
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1398 GAGCTCCAGAGTGTCTGTGTATACAGCTGTGTGTGTGTGTGTGTGT 1447
306 ValValLeuLysAlaArgHisLeuProLysSerAspValSerGlyLeu 323
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1448 GGTGTCTCAAGCCAGACACTGTGCAGAAAGATGATGTGCGGTCTCT 1497
323 er...AspProTyrValLysValAsnLeuTyrHisAlaLysArgIle 338
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1498 CAGTATTCCTTATGTACAGTGAACGTATACAGGACAGAAAGCATTT 1547
339 SerIlyLysIlyThrHisValLysLysCysThrProAsnAlaValPheAs 355
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1548 GCCAAGAAAGAAACCCATGTGAAAGAGTGTGCACTTGAACCCCATTTCA 1597
355 nGluLeuPheValPheAspIleProCysGluGlyLeuGluAspIleSerV 372
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1598 TGAATCTTTCATACACATCCCACTGACCTCTGCTGATATTCAGCA 1647
372 aGluPheValLeuAspSerGluArgLysSerArgAsnGluValIle 388
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1648 TCGAGTCTCTGTATGACATCTGTGTGTGTGTGTGTGTGTGTGTGT 1697
389 GlyIleuValLeuGlyAla...AlaIleGluGlyThrGlyGlyLysI 404
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1698 GGGAGGCTGATCTGGGGCCACACAGCTGTCAACGCTGTGTGTGAACA 1747
404 sTrpLysGluIleCysAspTyrProArgArgGlnIleAlaIleThrPheI 421
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1748 CTGAGAGAGGCTGTGCGAGAGCCCGGAGACCTGTGTGTGTGTGTGT 1797
421 allLeuCysAsp 424
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1798 GTCGTGAGCGAG 1808
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT.AAT29743
seq_documentation_block:
ID AAT29743 standard; cDNA to mRNA; 1876 BP.
XX
AC AAT29743;
XX
DT 29-APR-1997 (first entry)
XX
DE Mouse Inositol polyphosphate binding protein IP4-BP gene.
XX
KW Synaptogamin; antagonist; inhibitor; neurotransmitter; hormone;
KW calcium release; inositol; polyphosphate; pentakisphosphate;
KW hexakisphosphate; tetrakisphosphate; binding protein; IP4-BP;
KW PCR primer; polymerase chain reaction; murine; ds.
XX
OS Mus musculus.
XX
Key Location/Qualifiers
FH 16..1284
FT CDS
FT /tag= a
FT /product= IP4BP

```

XX JP08092290-A.  
 XX 09-APR-1996.  
 XX 20-SEP-1994: 94JP-0252942.  
 XX 20-SEP-1994: 94JP-0252942.  
 XX (SOSF-) SOSF KK.  
 XX WPI: 1996-236100/24.  
 XX P-PSDB: AAR97722.  
 DR Inositol polyphosphate (IPP)-binding peptide - useful as an  
 PT inhibitor of calcium, neurotransmitter release or hormone release  
 PT and for identifying IPP antagonists  
 XX  
 XX Example 8: Page 12-14: 19pp: Japanese.

CC Degenerate primers (see AAT29741 and AAT29742) were designed based on  
 CC peptide fragments from the inositol polyphosphate binding protein  
 CC IP4BP isolated from ddy mouse cerebellum. The primers were used for  
 CC PCR amplification of the mouse IP4BP gene (see AAT29743). The C2A  
 CC domain at the C-terminal end of IP4BP contains the synaptogramin II  
 CC peptide. A consensus sequence was obtained from a comparison of  
 CC synaptogramin peptides from a variety of species. Peptides comprising  
 CC the consensus motif (see AAR97721) are able to bind to inositol  
 CC tetrakisphosphate. Inositol pentakisphosphate and inositol  
 CC hexakisphosphate: such peptides can inhibit neurotransmitter release  
 CC and hormone release and are calcium inhibitors.  
 XX

Sequence 1876 BP; 481 A; 491 C; 504 G; 400 T; 0 other;

# Alignment scores:

Quality: 604.00 Length: 422  
 Ratio: 2.359 Gaps: 12  
 Percent Similarity: 60.664 Percent Identity: 36.256

## Alignment block:

US-09-680-121-2 x AAT29743 ..

Align seg 1/1 to: AAT29743 from: 1 to: 1876

```

10 GluPheAspGluIlePro.....ThrValaIglYIlePheSe 22
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
184 GAGATCAGACAGATGCGCCCTGCGGCTGATGCGCATGCGTGT 233
22 rIalPheGlyLeuValrPheThryaISeIleuPheAla.....TrpI 36
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
234 GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 283
36 IeCyScyGlnArgIySerSerIySerAsnIyThrProtyrIyLys 52
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
284 GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
53 rHeValnISeValIeUySgIyValaISeIleTyrrProGluAsnIeAsnSe 69
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
322 .....AAAGGCATG.....AAAGGCATGAAAT 347
69 rlySlyIySlyrPheGlyAlaAspAsnIySngIyValIySAsnIySProA 86
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
348 GAAAGACATGAAGGGGCGCAGATGATGATGCA..... 384
86 IalValrProLyAsnSerIeUyIleUyAspIeUyIyUyAspIeUyAsn 102
384 ..... 384
103 GlyAsnPheProLyThrAsnIeUySProGlySerProSerAspIeUy 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
385 .....GAGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 405

```

```

119 uAsnAlaThrProLySleuPheUyGluGlyIyUyGluSerValSerP 136
406 .....GAGGTGAAGGCCAG..... 420
136 roGluSerIeUySserThrSerIeUyThrSerIeUyIyUyGluGlu 152
421 .....GAGGAGAAAGACCAAG 438
153 rlySleUyIyThrIeUyPheSerIeUyIyUyAsnPheGlyUyArgIySAl 169
439 AACCTGGCCAAATTCGATTCCTGACATATGATTTCCAGGCCACCA 488
169 aPheValaIAsnIleUySgIyUyAlaArgIyLeuProAlaMetAspIyG 186
489 GCTCACCCTGGGTGCTGCTGACGCTGCGGAATCCAGCCCTGGAC...A 535
186 InSeMetThrSerAspProTyrrIleUyMetThrIleUyProGlyUyS 202
536 TGGGTGCACATCAGACCTTATGTCAAAGTCTTCCTGCTCCAGCAAG 585
203 rlySleUySValIySThrArgValIeUyArgIySThrIeUySProAlaP 219
586 AAGAGAAATATGACACTAGGTGATGAGAGAGCGTGAAACCCAGCTT 635
219 eAspGluIyPheThrPheTyrrIyIeProTyrrThrGlnIleGlnGlu 236
636 CAAATGACATTCACCTTC...AGGTCCATACACAGAGATTAGAGCA 682
236 eUalIeUySThrIleUySerPheAspArgPheSerArgAspAsp 252
683 AGACCTGGTATGAGCAATCTATGACTTGACCCCTCTCTAGCATGAC 732
253 IleIleGlyIyUyValIeUyIleUySerGlyIleUyIeUySerGlu 269
733 ATCATCGGGAGGTGAGGTACCATGACACAGTGACCT.....G 776
269 yIySmetLeu.MetAsnArgGluIleIleUySArgAsnValaArgIyS 285
777 CAGACCATCGAGAGAAATGAGACCTACAAAGCGGAGAGAAAGAGAC 826
286 SerGlyArg..GlyIeUyLeuIleSerIeUyCyTyrrGlnSerThr 301
827 CAGAGAAATGGGTGACATCTGTACCTCTGCGCTGACCTGACACAGCT 876
302 AsnThrIeUyThrValaIyValIeUySAlaArgIySleuProLySers 318
877 GCGAGCTGACCGCTGTATCTGCTGAGGCCAAGAACTGAAAGATGCA 926
318 rValSerGlyUySerAspProTyrrValIyValaIAsnIeUyTrnIAla 335
927 CGTAGGGGCTTTCAGACCCCTATATGTAAGATCAGCTGATGACAGAG 976
335 yIySValrIleSerIySlySlyThrIySValIySlyScyThrProAsn 351
977 GTAAAGACATCAAGAAAGAAAGAGACAGTGAAGAAAGAACCCCTGAA 1026
352 AlalValrPheAsnGluUyPheValrPheAspIleProCySgIyUyLeu 368
1027 CCGTACTCAAGAGCTTCAGCTTCAGATCCCTTTTGAAGAGATCA 1076
368 uAspIleSerValIyUyPheUyValIeUyAspSerGluArgIySerArg 385
1077 GAAAGTTCAGAGTGGTGGTCCAGCTGCTGATGACAAACTGGGAGAA 1126
385 sngIyUyAlIleGlyIyUyValIeUyGlyAlaIalalagIyUyTrgIy 401
1127 ATGAAGCCATCGAAAGATCTTGTAGGACACAAAGCCAGCAGCCAG 1176
402 GlyUyIySThrIySgIyUyIleUySAspTyrrProArgArgIyIleAla 418
1177 TTGCGGCACTGGTGGACATGCTGGSCAAACCTGGAGGCCCTTTGCCA 1226
418 sTrnHisValIeU 422

```





```

358 evalPhenaspIleProcySGluGlyLeuGIuaspIIeserValGIuPheU 375
      | ||| ...||| ||| |||||...: |||...: |||
1335 CAGCTTCAAAgTCCCCAGAGAAgACGGAATAATGCCAGCTAgTGTTA 1384
      |||||: ...: ||| |||||: |||||: |||||:
375 euValIeAAsperGIuArgGlySerTrgAsnGIuAlIegiGlyGIuLeu 391
      |||||: ...: ||| |||||: |||||: |||||:
1365 CAGTTTGGGCCACAACTGAAGAGCAGCAATgACTTTCATGGGAGAGATC 1434
      |||||: |||||: |||||: |||||: |||||: |||||:
392 ValIeGIuAlaIaAlaGIuGIu...ThrGIyGIuIuHISTrPlySGI 407
      |||||: |||||: |||||: |||||: |||||: |||||:
1435 GTCATGTGGCCAGTACTCTTCAGGCCCTCTGAGACCAACACACTGGAGGGC 1484
      |||||: |||||: |||||: |||||: |||||: |||||:
407 uIleCysAsPTyrProArgArgGIuIleAlaLySTrPHisValIeU... 422
      ::::: ::::: ||| ::::: ||||| |||
1485 CATGCTCAACACGACCGCACCGCGTGAGAGAGTGCATAgCCTGAGGT 1534
      423 ..... CysAsp 424
      1535 CCCGAGCTGAGTGTgAC 1551
seq_name: /SID58/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI61195
seq_documentation_block:
ID AAI61195 standard; cDNA; 1569 BP.
XX
XX AAI61195:
XX
XX 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 5184.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Dimmac RT;
XX
XX WPI: 2001-442253/47.
XX
XX P-PSDB; AAM42039.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 5184; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX

```





XX Homo sapiens. Location/Qualifiers  
 XX Key 21-2135 /tag=a  
 XX CDS /product= Rabphilin-3A  
 XX  
 XX JP06184199-A.  
 XX  
 XX 05-JUL-1994.  
 XX PD  
 XX 24-DEC-1992: 92JP-0344055.  
 XX  
 XX 24-DEC-1992: 92JP-0344055.  
 XX  
 XX (EISA ) EISA CO LTD.  
 XX  
 XX WPI: 1994-252836/31.  
 XX P-ESDB; AKR57421.  
 XX  
 XX Target protein of a low molecular G protein rabphilin-3A (RAB3A)  
 XX - found in the brain and involved in release of nerve transmitter  
 XX substance  
 XX  
 XX Claim 1: Page 6-9; 9pp: Japanese.  
 XX  
 XX This sequence encodes a low molecular weight G protein target  
 XX protein designated rab3a p25. Rabphilin-3A (Rab3A) is distributed  
 XX specifically in brain tissue and participates in the release of  
 XX nerve transmitter substance and is useful in the study of its  
 XX secretion.  
 XX  
 XX Sequence 2135 BP; 552 A; 609 C; 630 G; 344 T; 0 other;

alignment\_scores:  
 Quality: 408.00 Length: 337  
 Ratio: 1.934 Gaps: 4  
 Percent Similarity: 62.611 Percent Identity: 29.377

alignment\_block:  
 US-09-680-121-2 x AA067062 ..

Align seg 1/1 to: AA067062 from: 1 to: 2135

112 ProlyserProseraspleuGlusAlaThrProlysleupheleuGl 128  
 1101 CCGGGCTCTATACCAAGCTCTGCGGCTGCCACCGCGGTGGC 1150  
 128 uGlyGluLysGluSerValSerProGluSerLeuLysSerThrSerL 145  
 1151 CTCGGCCCTCAGCCGCCCTGAGAGAGAGAGAGAGAGAGAGAG 1200  
 145 eutHrserGluLysGluLysGluLysLeuGlyThrLeuPhePheSerL 161  
 1201 GCTACGATTCGATGAACACACACCTGGGTGCTGAGTTCACTT 1250  
 162 GluTrAsnPhenGluArgLysAlaPheValAlaAsnLeuGluAla 178  
 1251 CTGTATGACCAAGCAACAGCTCCGACACCAACCAACCAACCA 1300  
 178 gGlyLeuProAlaMetAspGluLysSerMetThrSerAspProTyr 195  
 1301 GGAGCTGAAGCCCATGATTCGAAATGGCTTG...GCCGATCCCT 1347  
 195 ysmetThrIleLeuPro.....GluLysLysLysLysValLysThr 209  
 1348 AGCTGCACCTCTCTCCGCGAGCCAGCAAGTCACACACCTTCG 1397  
 210 ValLeuArgLysThrLeuAspProAlaPheAspGluThrPheThr 226  
 1398 ACCCTGCGGACACCAACCAACCAACCAACCAACCAACCAACCA 1447

226 rGlyIleProTyrThrGluIleGluLysAlaLeuHisPheThrIle 243  
 1448 TGGCATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497  
 243 euserPheAspArgPheSerArgAspAspIleIleGlyLysVal 259  
 1498 GTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1547  
 260 ProLeuSerGlyIleGluLeuSerGlu..... 268  
 1548 TCCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1597  
 269 .....G 269  
 1598 GGAGCGGTGATCCGATGAGAGAGAGAGAGAGAGAGAGAG 1647  
 269 LysMetLeuMetAsnArgGluIleIleLysArgAsnValArgLys 285  
 1648 GCATGGCCCTCTATGAGAGAGAGAGAGAGAGAGAGAGAG 1694  
 286 SerGlyArgGlyGluLeuLeuIleSerLeuCysTyrGlnSerThr 302  
 1695 GAGGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1744  
 302 nThrLeuThrValValLeuLysAlaArgHisLeuProLysSerAsp 319  
 1745 AGGCGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1794  
 319 alserGlyLeuSerAspProTyrValLysValAsnLeuThrHis 335  
 1795 CCATGGCTCTCATGAGAGAGAGAGAGAGAGAGAGAGAG 1844  
 336 LysArgIleSerLysLysThrHisValLysCysThrProAspAl 352  
 1845 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1894  
 352 aValPheAsnGluLeuPheValPheAspIleProCysGluGly 369  
 1895 TGAATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1944  
 369 spLysSerValGluPheLeuValLeuAspSerGluArgLysSer 385  
 1945 AGAGTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1994  
 386 GluValIleGlyLysLeuValLysGlyAlaAlaIleGluThrGly 402  
 1995 GATTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2044  
 402 yGluHisTrpLysGluIleCysAspTyrProArgArgGlnIle 419  
 2045 AAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2094  
 419 rPHISValLeu 422  
 2095 GGACACAGCTA 2105  
 seq\_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:AA00321  
 seq\_documentation\_block:  
 ID AA00321 standard; CDNA to mRNA; 2255 BP.  
 AC AA00321;  
 XX  
 XX 22-OCT-1999 (first entry)  
 XX  
 XX Mouse Doc2alpha gene sequence.  
 XX  
 XX Transgenic mouse; Doc2alpha; nervous disease; endocrine system; ds.  
 XX  
 XX Mus sp.  
 XX  
 XX key Location/Qualifiers  
 FH







```

264 1eg1u1eSerg1uGly1ys.....Met1ueWetAsnArg1u1le 277
      ::::: |||::: |||
726 TCAGACCTTCGACAGAAAGCATTTTAACATCTGCCTCGAGGCCCAAGTC 775
277 ..... 277

776 CCGCTGGGGCTCCCCCTTCATCATCTCAGCGGGCGGTAGAGGGCATCTCC 825
278 ...11e1ysArgAsnValArg1ysSerSerg1y..... 287
      ::|||
      :::::||||
826 TTATCTGAAGGACCTTCGACGACGGCGAGCAGGGCAGGGCTGCTGAGAG 875
288 ..Arg1yG1u1e1u1e1eSer1euc1y1yG1nSert1rr1Asn1Thr 303
      :::::|||||:::||||| |||:::||||
876 ACGCGGGCGGACATCCTGCTGAGTTCAGTACACCTTCGCGCGCGCGGA 925
304 1eu1rr1Val1Val1e1u1e1yAlArG1n1S1e1uProl1ysSer1Asp1Se 320
      ||| ||| :::::|||||
926 CTGCTGCTGAGGACATCTTCGCTGCGGCCCATCTGAGCTCCATGAGCTCA 975
320 rG1y1e1uSer1AspProl1yVal1ysVal1Asn1e1u1Thr1n1Al1ys1ys 337
      ::||| ||||| ||||| |||||::: |||
976 CGGTACTACGACCCCTACGCTCAGACAGATACCTGACGGCCCGATGTGGACA 1025
337 rG11Se1r1ys1ys1yThr1n1S1Val1ys1ysC1y1Thr1Pro1Asn1AlVal 353
      ::::: ||||| ||||| ||||| ||||| |||
1026 ACAAATCCAGCATAGACGTGTGTGAAGAAAGAACTCTCAACCCAGAA 1075
354 PheAsnG1u1e1uPheVal1PheAsp1r1eP1oc1ySg1uG1y1e1uG1uAsp1r 370
      ||||| ||| ::::: |||
1076 TTTAACGAGAGACTTTTCTTACGAGATAGAGTCTCCACTCTGGCCACCAA 1125
370 eSer1ValG1uPhe1e1uVal1e1uAspSerG1uArg1ySer1ArgAsnG1uV 387
      ::::: ||| ||| ::::: |||::: |||:::
1126 GACCTCGAAGACACCGCTGCGACATAGACATTGGCAATCCAAATGACT 1175
387 a1l1eG1y1u1e1uVal1e1uG1y1AlAlAlAaG1uG1y1ThrG1yG1yG1u 403
      ||||| ::::: ||||| ::::: ||||| ::::: |||
1176 TCATTGCGGGCGCTGCTCCTGGGGCCAGGTGCCGAGGCGAGGCTCGGAG 1225
404 H1Str1P1ySg1u1leC1yAsp1Ty1rPro1Arg1n1leAl1ay1Str1P1H 420
      |||||::: |||::: |||::: |||::: |||
1226 CACTGAGTGAAGTGCCTGCACAGCAGCGGAGCAGCACCCTGAGCGCTGGCA 1275
420 sVal1e1u 422
      |::: |||
1276 CACCTCG 1282

seq_name: /SID8/gcgdata/geneseq/geneseqn/NA2001.DAT.AA110603
seq_documentation_block:
ID AA110603 standard; DNA; 383 BP.
XX
XX AA110603:
XX
XX 12-OCT-2001 (first entry)
DT
XX
XX Probe #536 for gene expression analysis in human cervical cell sample-
KW Probe: human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
XX Homo sapiens.
OS
XX MO200157278-A2.
XX
XX PD
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001MO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX

```

30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0532366.  
 PR 21-SEP-2000; 2000US-0534687.  
 PR 27-SEP-2000; 2000US-0536359.  
 PR 04-OCT-2000; 2000GB-0024283.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PA Penn SG, Hanzel DK, Chen W, Rank DR:  
 XX  
 DI WPI: 2001-488901/53.  
 XX  
 XX Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 25; SEQ ID NO 536; 487bp; English.  
 CC  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENP). The present sequence is one such probe. The SENPs are derived  
 CC from human head cells. The SENPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes described are useful in grading and/or staging  
 CC of diseases of the cervix notably cervical intraepithelial neoplasia  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SO Sequence 383 BP; 125 A; 71 C; 58 G; 123 T; 6 other;

alignment\_scores:  
 Quality: 370.50 Length: 88  
 Ratio: 4.464 Gaps: 3  
 Percent Similarity: 94.318 Percent Identity: 90.909

alignment\_block:  
 US-09-680-121-2 x AM10603/rev ..

Align seg 1/1 to reverse of: AM10603 from: 1 to: 383

```

197 ThrlleuPProglulysHisValysThrArgValleuArgly 213
|||||
381 ACGATCTCCAGAGAGACATTAAGTGAAGCTAGACTGAGAAA 332
|||||
213 sThrleuAspProAlaPheAspGluThrPheThrPheTyrglyleProT 230
|||||
331 AACCTTGATCCAGCTTTGATGAGACCTTACATTCATGAGATACCT 282
|||||
230 yThrGlnIle..GlnGluLeuAlaLeuHisPheThr.IleLeuSerPh 245
|||||
281 ACACCCAAATCCNNAAGATTGCNCTTCGACCTTCACAAATTTTGAGTT 232
|||||
245 eaAPArpPheSerArg.AspAspIleIleGlyGluValleuIleProLeu 261
|||||
231 TGACAGCTTTCAAGANAGATGATCATTCGGAGCTTCATATTCCTCTC 182
|||||
262 SerGlyIleGluLeuSerGlyGlyMetLeuMetAsnArgGluIleI 278
|||||
181 TCGGAAATTCATTTATCTGAAGAAAAATGTTAATGATAGAGATCA 132
|||||
278 elysArgAsn 281
|||||
131 CAAGAGAAAT 122

```

OM of: US-09-680-121-2 to: EST:\* out\_format : pfs  
Date: Dec 12, 2001 11:59 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

```
-MODEL=fmet+p2n.model -DEV=xlp  
-O=/cgn2_1/USPTO.spool/US09680121/runat_12122001_105511_22723/app_query.fasta.1.488  
-DB=EST -QMT=fastap -SUFFIX=p2n.rst -GAP=12.000 -GAPEXT=4.000  
-MIMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAP=4.500  
-OGAPEXT=0.050 -XGAP=10.000 -XGAPEXT=0.500 -FGAP=6.000  
-FGAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500 -DELCP=6.000  
-DELETE=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCLINK=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=2000000000 -USER=US09680121 -CGEN1_1_0 -NCPU=6 -ICPU=3  
-LONGLOG -NO_XLPHY -WAIT -THREADS=21
```

## Search information block:

Query: US-09-680-121-2

Query length: 425

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 1537.320000

## score list:

Sequence	Strnd	Orig	Zscore	Escore	len	Documentation
gb_est1.AM160503	+	1065.00	2047.99	6.6e-105	656	AM160503 au73c12.y1 Schneider
gb_est1.AM160503	+	1052.00	1923.08	1.6e-103	635	AM160503 au73c12.y1 Schneider
gb_est1.BG655092	+	1016.50	1952.71	1.3e-99	739	BG655092 DRABY001 Rat DRG Libra
gb_est1.AV066332	+	845.00	1632.52	3.3e-81	565	AV066332 AVE66332 Bos taurus k1
gb_est1.BG261870	+	802.50	1532.90	3.2e-76	1174	BG261870 602373670F1 NIH.MGC.9
gb_est1.BG742430	+	746.00	1431.16	1.5e-70	533	BG742430 DRNC106 Rat DRG Libra
gb_est1.BG744033	+	711.50	1359.37	1.5e-66	888	BG744033 602722813F1 NIH.MGC.10
gb_est1.BF937960	+	700.00	1339.46	1.9e-65	671	BF937960 fm10905.y1 Zebrafish a
gb_est1.BG671533	+	688.00	1314.56	4.7e-64	791	BG671533 DRNB005 Rat DRG Libra
gb_est1.BF402440	+	687.00	1318.49	2.8e-64	434	BF402440 UI-R-CA0-bhp-c-07-0-UI
gb_est1.AW367632	+	668.00	1278.63	4.7e-62	589	AW367632 MR0-HR0165-121199-003-
gb_est1.AW367662	+	668.00	1278.63	4.7e-62	589	AW367662 MR0-HR0165-191199-003-
gb_est1.AM914162	+	665.00	1271.28	1.2e-61	689	AM914162 EST344466 Normalized
gb_est1.AM914163	+	665.00	1270.96	1.3e-61	712	AM914163 EST344467 Normalized
gb_est1.AW035352	+	639.50	1220.17	8.4e-59	813	AW035352 AU035352 Suga mouse
gb_est1.AA081755	+	632.50	1211.00	2.7e-58	518	AA081755 zn2207.r1 Stratagene
gb_est1.BF563293	+	629.50	1204.48	6.3e-58	556	BF563293 UI-R-B01-aj1-e-05-0-UI
gb_est1.BF044050	+	626.00	1196.93	1.7e-57	601	BF044050 BP250004A20C6 Soares R
gb_est1.BF043853	+	592.00	1131.48	7.3e-54	569	BF043853 BP250004A20C6 Soares R
gb_est1.BG667100	+	590.00	1127.99	1.2e-53	547	BG667100 DRACD03 Rat DRG Libra
gb_est1.AU050945	+	586.50	1117.98	4.2e-53	760	AU050945 AU050945 Suga mouse
gb_est1.BG666223	+	586.00	1123.19	2.1e-53	404	BG666223 DRABAE06 Rat DRG Libra
gb_est1.B1112436	+	557.50	1061.79	5.6e-50	753	B1112436 602900354F1 NCI.CGAP.J
gb_est1.AV728050	+	549.00	1051.53	2.1e-49	398	AV728050 AV728050 HTC Homo sapi
gb_est1.H35435	+	534.00	1033.27	2.2e-48	354	H35435 EST10556 Rat PC-12 cell
gb_est1.BG669842	+	527.00	1021.62	9.7e-48	432	BG669842 DRNAH10 Rat DRG Libra
gb_est1.BF472172	+	508.50	970.63	6.7e-45	504	BF472172 UI-M-BH3-awx-d-11-0-UI
gb_est1.BG084832	+	469.00	892.06	1.6e-40	613	BG084832 H310608-5 NIA Mouse 1
gb_gss.A2889578	-	469.00	892.06	1.6e-40	613	A2889578 RPI1-24-168BD.TV RPI1
gb_est1.AV663576	+	468.00	884.82	1.2e-40	399	AV663576 B663576 Bos taurus r1
gb_est1.A520824	+	467.00	884.80	1.0e-40	866	A520824 A520824 LTI.NFL004.NE
gb_est1.AA445762	+	455.00	868.27	3.4e-39	434	AA445762 v63B02.s1 Knowles Sol
gb_est1.AV727109	+	455.00	859.96	1.0e-38	1054	AV727109 AV727109 HTC Homo sap
gb_est1.AL516750	+	451.00	863.31	6.4e-38	843	AL516750 AL516750 LTI.NFL011.NE
gb_est1.BB254363	+	451.00	863.31	6.4e-38	336	BB254363 BB254363 RIKEN full-16
gb_est1.BG667224	+	428.00	815.83	2.8e-36	436	BG667224 DRAB08 Rat DRG Libra
gb_est1.A1322673	-	428.00	815.83	2.8e-36	852	A1322673 m121f03.y1 Soares mous
gb_est1.BE947248	-	426.00	805.92	1.0e-35	808	BE947248 UI-M-BH3-awx-d-11-0-UI
gb_est1.BF918236	+	426.00	805.92	1.0e-35	808	BF918236 SMCV3MCA454E08SK Oncho
gb_est1.A1322332	+	419.00	790.41	7.3e-35	984	A1322332 mc27605.y1 Soares mous
gb_est1.B1411213	+	419.00	790.41	7.3e-35	984	B1411213 602964537F1 NCI.CGAP.J

gb\_est1.AM174543 + 418.50 794.45 4.4e-35 589 ! AM174543 f105d11.y1 Zebrafis  
gb\_gss.CNS036RD + 418.00 788.61 9.2e-35 652 ! AL230386 Tetradon nigroviri  
gb\_est1.A1878251 + 406.00 773.04 6.8e-34 655 ! A1878251 fc52d10.y1 Zebrafis  
gb\_gss.CNS03172 + 406.00 764.82 1.9e-33 1021 ! AL245240 Tetradon nigroviri

seq\_name: gb\_est1.AM160503

## seq documentation block:

LOCUS AM160503 656 bp mRNA EST 09-NOV-1999

DEFINITION au73c12.y1 Schneider, fetal brain 00004 Homo sapiens cDNA clone  
IMAGE:2781910 5' similar to SW:SYT4\_RAT STYAP232 SYNAPTOTAGMIN IV. ;  
mRNA sequence.

ACCESSION AM160503  
VERSION AM160503.1 GI:6299536

KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 656)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,  
Kizman, D., Kudada, T., Lacy, M., Le, N., Lennon, G., Maitra, M., Martin,  
J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theisling, B.,  
White, Y., Wylie, T., Waterston, R. and Wilson, R.

White, Y., Wylie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST Project  
Unpublished (1997)

Other ESTs: au73c12.x1  
Contact: Wilson RK

Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@wustl.edu  
This clone is available royalty-free through LINL: contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40RP from Gibco

High quality sequence score: 427.  
Location/Qualifiers

1..656  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2781910"  
/clone\_lib="Schneider fetal brain 00004"  
/sex="male"  
/tissue\_type="frontal lobe"  
/dev\_stage="5 months post-conception"  
/lab\_host="DH10B"

/note="Organ: brain; Vector: pBluescript SK (Stratagene);  
Site\_1: SstI; Site\_2: XhoI; Double-stranded cDNA was  
prepared from human fetal brain tissue. 5' and 3' adaptors  
were used in cloning as follows: 5' adaptor  
sequence: 5'-GAGAGAGAGAGAGCTCAAGATCTTAATTAATTAATTCCTCCCTCCCTCC-3'  
and 3' adaptor sequence:  
5'-GAGAGAGAGAGCTCAAGATCTTAATTAATTAATTCCTCCCTCCCTCC-3'

size-selected for >0.5 kb inserts and has an average  
insert size estimated at 1.2 kb. This library was  
constructed using the CAP-trapper method for full-length  
enrichment and has not undergone amplification. Library  
was constructed by Dr. Claudio Schneider (LINCB-Area  
Science Park, Trieste, Italy)."

BASE COUNT 213 a 147 c 138 g 156 t 2 others  
ORIGIN

alignment\_scores: Quality: 1065.00 Length: 214  
Ratio: 5.071 Gaps: 0  
Percent Similarity: 98.131 Percent Identity: 97.196

alignment\_block:  
US-09-680-121-2 x AM160503 ..



265 AGCCAGCTGTGCCAAGAATTCATTCATCTGGATCTTGAAAGAGAGAT 314  
 101 LeuAsnGlyAsnProlySerThrAsnLeuLysProGlySerProSerAs 117  
 315 CTCGAAGGCAATTTTCCCAAAACCAACCTCAACCTGGCAGTCTCTG 364  
 117 pleuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGlySerV 134  
 365 TCTGGAGATGCAACCCGAGCTCTTTAGAGGGGAAAAAGATGAC 414  
 134 aSerProGlySerLeuLysSerSerThrSerLeuThrSerGluLys 150  
 415 TTTCCCTGAGAGTTTAAAGTCACGACTTCCCTTACTTCCAGAGAAA 464  
 151 GlnGluLysLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGlu 167  
 465 CAGAGAACCTGGGAACCTCTCTCTCTCTGAGATACAACTTCAGAG 514  
 167 GlyAlaPheValAlaAsnIleLysGluAlaArgGlyLeuProAlaMet 184  
 515 AAAGCATTTTGTCATATTCAGAGAGCCCTGGCTTCCAGCCCTTG 564  
 184 spGluInSerMetThrSerAspProTyrIleLysMetThrIleLeuPro 200  
 565 ATGAGCAGTCGATGACCTCTGACCCATATATCAAAATGACGATCTCC 614  
 201 GluLysLysHisLysValLys 207  
 615 GAGAGAGAGCTTAAAGTAA 635

seq\_name: gb\_est2:BG665092

seq\_documentation block:

LOCUS BG665092 739 bp mRNA EST 30-APR-2001  
 DEFINITION DRABYCO1 Rat DRG Library Rattus norvegicus cDNA clone DRABYCO1 5',  
 mRNA sequence.

ACCESSION BG665092  
 VERSION BG665092.1 GI:13887014

KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 739)  
 Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G.,  
 Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z. and  
 Zhang,X.  
 Distinct gene expression profiles of rat dorsal root ganglion  
 induced by peripheral nerve axotomy  
 Unpublished (2001)

TITLE  
 JOURNAL  
 COMMENT  
 Contact: Zhang Xu  
 Laboratory of Sensory System  
 Institute of Neuroscience  
 320 Yue Yang Road, Shanghai 200031, P.R.China  
 Tel: 86-21-64748700-121  
 Fax: 86-21-64713446  
 Email: xu.zhang@ion.ac.cn

This clone is also available at Chinese National Human Genome  
 Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,  
 Pudong New Area, P.R.China. Please contact with Zhang Xu  
 (xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@hgc.sh.cn)  
 PCR primers  
 FORWARD: T3  
 BACKWARD: T7  
 Seq primer: T3  
 POLYA-No.

FEATURES  
 source

1. 739  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="DRABYCO1"

/clone\_lib="Rat DRG Library"  
 /sex="male"  
 /tissue\_type="dorsal root ganglion"  
 /dev\_stage="adult"  
 BASE COUNT 210 a 196 c 168 g 165 t  
 ORIGIN  
 alignment\_scores:  
 Quality: 1016.50 Length: 247  
 Ratio: 4.458 Gaps: 2  
 Percent Similarity: 92.308 Percent Identity: 83.806  
 alignment\_block:  
 US-09-680-121-2 x BG665092 ..  
 Align seg 1/1 to: BG665092 from: 1 to: 739  
 4 ILeThrSerArGluGluPheAspGluIleProThrValAlaGlyI 20  
 1 ATCACCAACGCGCCGCGGAAATTCGATGAATTCACACAGTGGCAT 50  
 20 ePheSerAlaPheGlyLeuValPheThrValSerLeuPheAlaTrpIleC 37  
 51 CTTCAGTGTCTTTGGCTGTCTTCACTGTCTCTCTTGGCTGATCT 100  
 37 yScySGlnArGlySer.SerLysSerAsnLysThrProTyrLysPh 53  
 101 GCTGCCAGAGAGATCAGCTAAGTCCACACAGATCTCTCCACATCAAGATT 150  
 53 eValHisValLeuLysGlyValAspIleTyrProGluAsnLeuAsnSerL 70  
 151 TGTGCACGTCTTAAGAGATTGATATATCCAGAGAAACCTAATAGCA 200  
 70 yLysLysPheGlyAlaAspAspLysAsnGluValLysAsnLysProAla 86  
 201 AAAAGAGATTGGAGAGATGACAGACAGTGAACCAAGCTTAAGCAGCT 250  
 87 ValProLysAsnSerLeuHisLeuAspLeuGluLysArgAspLeuAsnG 103  
 251 TTGCTTAACCTTTCCTCGACCTGATCTGAGAGAGGAGACCTCAATGG 300  
 103 yAsnPheProLysThrAsnLeuLysProGlySerProSerAspLeuGlu 120  
 301 CAATTTCCCAAAACCAACCCCAAGCTGCAGCTCTTGATCTGAGAAA 350  
 120 snAlaThrProLysLeuPheLeuGluGlyLysGlySerValSerPro 136  
 351 ATGTACCCCAAAAGCTCTTCCGGAGACGGAAAAAGAGGCGCTCCCT 400  
 137 GluSerLeuLysSerSerThrSerLeuThrSerGluLysGlnGly 153  
 401 GAGAGCTTGAATCCACGACTTCCCTCATTCAAGAGAGAAACAGAGAA 450  
 153 sLeuGlyThrLeuPhePheSerLeuGluTyrAsnPheGluArgLysAla 170  
 451 GCTGGGCATCTCTCTTCTCTCTCAGAGTACAACTTGAGAAAGAGACAT 500  
 170 heValValAlaAsnIleLysGluAlaArgGlyLeuProAlaMetAspGlu 186  
 501 TTGGTGGACATCAAGAGAGCCAGGCGTTACACGCGATGATGAGACAC 550  
 187 SerMetThrSerAspProTyrIleLysMetThrIleLeuProGluLys 203  
 551 TCCATGACCTGTGACCCATCAATCAAGATGACAGATTCTTACAGAGAGAA 600  
 203 sHisLysValLysThrArgValLeuArgLysThrLeuAspProAlaPhe 220  
 601 GCACACAGTGAACCAACACAGTGGTGAAGACCGTGGACCCCTGTTG 650  
 220 spGluThrPheThrPheTyrGlyLysProTyrThrGlnIleGlnGluLeu 236  
 651 ACGAACCTTCACATCTCTATGGGTGCTTATCC.CACATCCAAGAGCTG 699

237 ALALEuHisPheThrIleLeuSerPheaspArgPheSer 249  
:::|||||  
700 TCCTGC..ACTCACAGCGTGAGTTTGCACAGTTTCA 735

seq\_name: gb\_est1:AV606332

LOCUS	AV606332	565 bp	MRNA	EST	30-AUG-2000
DEFINITION	AV606332 Bos taurus kidney fetus Bos taurus cDNA clone E1K1034B06 5' , mRNA sequence.				

AUTHORS	Sugimoto, Y., Hirotsune, S., Takasuga, A., Itoh, R., Jichonzono, A. and Suzuki, H.
TITLE	bovine CDNA sequencing
JOURNAL	Unpublished (2000)
COMMENT	Contact: Yoshikazu Sugimoto

seq\_name: gb\_est2:BG261870

```

422 ucysaspGly 425
      |||||
512 CTGTGATGGT 521

```

LOCUS	BG261870	1174 bp	mRNA	EST	13-FEB-2001
DEFINITION	6023737670F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4481416 5' mRNA sequence.				

Plate: LLAM10316 row: k. column: 17  
High quality sequence stop: 689.

	a	c	g	t	f
BASE COUNT	348	327	279	219	1
ORIGIN					others



## alignment\_scores:

Quality: 802.50 Length: 233  
Ratio: 3.993 Gaps: 7  
Percent Similarity: 86.266 Percent Identity: 75.107

## alignment\_block:

US-09-680-121-2 x BG261870 ..

Align seg 1/1 to: BG261870 from: 1 to: 1174

```

1 MetAlaProIleThrThrSerArgIleGluPheAspIleIleProThrVa 17
168 ATGCTCTCTATCACACACACCGCGTGAATTCGATGAATTCACACAGT 217
17 ValGlyIlePheSerAlaPheGlyLeuValPheThrValSerLeuPhe 34
218 GGGGGGCACTTCAGTGGCTTTGGCTTCGCTTCTCTCTCTTTG 267
34 IatPrlIecysGlnArgLysSerSerLysSerAsnLysThrProPro 50
268 CCGGATCTGCTCTCAGAGAGATCAGCCAAATCCAGACAGACTCTTCCA 317
51 TyrLysPheValHisValLeuLysGlyValAspIleTyrProGluAsn 67
318 TACAGTTTGTGCACGTCGTTAAAGAGTGTGATATCTACCCAGAAACCT 367
67 uAnSerLysLysPheGlyAlaAspLysAsnGluValLysAsn 84
368 AAGTACCAAAAGAGTTGGAGAGATGCAAGAGTGAAGTGAAGGTA 417
84 ySProlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
418 AACCCGCTCGCCCAACCTTCCCTGCATCTTGACCTAGAGAACGCGAGC 467
101 LeuAsnGlyAsnDheProLysThrAsnLeuLysProGlySerProSer 117
468 CTCAAATGGCAACTCCCAAGCAACCCAAAGCTGGCAGCTTCTCTGA 517
117 PleuGluAsnAlaThrProLysLeuPheLeuGluGlyGluGluSer 134
518 TCTGGAAATGTCACCCCAAGCTCTTACGAGACAGACAGAGAGGCCA 567
134 alSerProGluSerLeuLysSerSerThrSerLeuThrSerGlu.Glu 150
568 ATTCCTCGAGACTTGAATCCAGACACTTCCCTCAGCTCAGAGAGGAA 617
150 sGlnGluLysLeuGlyThrLeuPhePheSerLeuGlu...TyrAsnDhe 166
618 ACAAGAGAACTGGGAGACTTCTGTCTCTAGAAAGTACAACACTTCGG 667
166 LuArgLysAlaPheValAlaAsn.IleLysGluAlaArgLysLeuPro 182
668 AGAAGAAAGCATTTGGTGTGACACCTCAAGAACCCAGAGGCTTACAC 717
182 aMetAspGluGlnSerMetThrSerAspProTyrIleLysMetThr.Ile 198
718 CATGGAGAAAGCATTCATGAGCCCTGACACCTCAATCAATGACGATC 767
199 LeuProGluLysLysHisLys.....ValLysThrArg.V 210
768 TTTACAGAGAGAAAGCCCAAGGAGACCAAGGGCCCGAAGAAAGACACA 817
210 alLeuArgLysThrLeuAspProAlaPheAspGluThrPhe 223
818 CGTTTGAGAAACCGTG.....GAACATTT 843

seq_name: gb_est2:BG672430

seq_documentation_block:
LOCUS BG672430 523 bp mRNA EST 30-APR-2001
DEFINITION DRNCIC06 Rat DRG Library Rattus norvegicus cDNA clone DRNCIC06 5',
mRNA sequence.
```

```

ACCESSION BG672430
VERSION BG672430.1 GI:13894529
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 523)
AUTHORS Xiao,H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G.,
          Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z. and
          Zhang,X.
TITLE Distinct gene expression profiles of rat dorsal root ganglion
JOURNAL induced by peripheral nerve axotomy
COMMENT Unpublished (2001)
          Contact: Zhang Xu
          Laboratory of Sensory System
          Institute of Neuroscience
          320 Yue Yang Road, Shanghai 200031, P.R.China
          Tel: 86-21-64748700-121
          Fax: 86-21-64713446
          Email: xu.zhang@ion.ac.cn
          This clone is also available at Chinese National Human Genome
          Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
          Pudong New Area, P.R.China. Please contact with Zhang Xu
          (xu.zhang@ion.ac.cn) or Han Zeguang (hanzeg@nc.sh.cn)
PCR Primers
FORWARD: T3
BACKWARD: T7
Seq primer: T3
POLYA-No.
FEATURES
source location/Qualifiers
1..523
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="DRNCIC06"
/clone_1ib="Rat DRG library"
/sex="male"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
BASE COUNT 158 a 130 c 122 g 113 t
ORIGIN
```

```

alignment_scores:
quality: 746.00 length: 173
ratio: 4.605 gaps: 0
Percent Similarity: 93.642 Percent Identity: 86.127
```

## alignment\_block:

US-09-680-121-2 x BG672430 ..

Align seg 1/1 to: BG672430 from: 1 to: 523

```

108 ThrAsnLeuLysProGlySerProSerAspLeuGluAsnAlaThrPro 124
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3 ACCAACCCCAAGCTGGACGCTCTTGATCTGGAATGTCACCCAAA 52
124 sLeuPheLeuGluGlyLysGluSerValSerProGluSerLeuLys 141
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
53 GCCTCTTCGGAGAGCGAAAGAGCGCGTCTTCCGACAGCTTGAGT 102
141 erSerThrSerLeuThrSerGluGluLysGlnGluLysLeuGlyThrLeu 157
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
103 CCAAGCACTTCCCTCATTCAAGAGAGAAACAAGAGAGCTGGCCACTC 152
158 PhePheSerLeuGluLysThrAsnDheGluArgLysAlaPheValAla 174
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
153 TTTCTGTCTAGAGTCAACTTCGAGAGAAAGCATTTGGGGAACAT 202
174 elysGluAlaArgLysLeuProAlaMetAspGluGlnSerMetThrSera 191
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
```

203 CAAAGAGCCCGAGGCTTACCAGCCATGATGACAAATCCAGACTCTG 252  
 191 sPpRtYrTlLeuSMeThrTlLeuProGluLysHISLysValLys 207  
 |||||  
 253 ACCCAATACATCAAGATGACATCTTACCAGAAAGAACCAAGAGAA 302  
 208 rThArgValLeuArgLysThrLeuAspProAlaPheAspGluThrPheTh 224  
 |||||  
 303 ACCAGAGTGTCTAGAGAACGCTGACCCGCTGTTGAGAAACCTTCAC 352  
 224 rPheTyrgLyLLeuProTyThrGlnIleGlnGluLeuAlaLeuHisPheT 241  
 |||||  
 353 ATTCTATGGGCTCCTTATGCCCATTTCAAGAGCTGTGCTGCACTTCA 402  
 241 hTlLeuSerPheAspArgPheSerArgAspIleIleGlyGluVal 257  
 |||||  
 403 CCGGCTGTGTTTGACAGTTTCAAGAGATGATGTCATTTGAGAGATC 452  
 258 LeuIleProLeuSerGlyIleGluLeuSerGlyLysMetLeuMetAs 274  
 |||||  
 453 CTCGTTCCCTTTCAGAGATTGATTTCTCAGACGGAAATGTTAATGAC 502  
 274 nArgGluIleIleLysArg 280  
 |||||  
 503 CAGAGATCATCAAGAGA 521

seq\_name: gb\_est2:BG744033

seq\_documentation\_block: 858 bp mRNA EST 15-MAY-2001  
 LOCUS BG744033 602722813f1 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:4849251 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG744033  
 VERSION BG744033.1 GI:14054686  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 858)  
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs.femail.nih.gov  
 Tissue Procurement: Dr Daniel McVicar, DBS/MCI  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM1688 row: j column: 04  
 High quality sequence stop: 792.  
 Location/Qualifiers  
 1. 858  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4849251"  
 /tissue\_lib="NIH\_MGC\_106"  
 /tissue\_type="natural killer cells, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="organ: blood; Vector: pOTB7; Site1: XhoI; Site2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACAGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH-MGC library."

BASE COUNT 230 a 231 c 213 g 184 t  
 ORIGIN

# alignment\_scores:

Quality: 711.50 Length: 244  
 Ratio: 3.471 Gaps: 3  
 Percent Similarity: 84.016 Percent Identity: 61.066

## alignment\_block:

US-09-680-121-2 x BG744033

Align seg 1/1 to: BG744033 from: 1 to: 858

116 SerAspLeuGluAsnAlaThrProLysLeuPheLeuGluGlyLysGlu 132  
 |||  
 81 AGCCCTATTACAAAGCTGACCCCT .....GGGAGAGCAA 115  
 132 uSerValSerProGluSerLeuLysSerSerThrSerLeuThrSerGlu 149  
 |||||  
 116 AACCACTCTCA .....TCATCTCAG 138  
 149 LuLysGlnGluLysLeuGlyThrLeuPhePheSerLeuGluThrAspHe 165  
 |||||  
 139 AGGAGAGATGATGCTAGATCCCTCACCCTTCTCAGTGGACTATTAACCTT 188  
 166 GluArgLysAlaPheValAlaAsnIleLysGluAlaArgGlyLeuProAl 182  
 |||||  
 189 CCGAAAAAGCCCTGTGTGACAAATCCAGAGCCCAAGCGCTGCGCAGT 238  
 182 aMetAspGluGlnSerMetThrSerAspProTyrlLeuSMeThrTlLeu 199  
 |||||  
 239 GATGATAGACAGACAGCCAGGA.TCTGACCCCTACATCAAAATGACCATTC 287  
 199 euPProGluLysHISLysValLysThrArgValLeuArgLysThrLeu 215  
 |||||  
 288 TTCTCGCAAAACGCGCATCGGTGAAGACAGAGCTGCTCGGAGAACCTTG 337  
 216 AspProAlaPheAspGluThrPheThrPheTyrgLyLLeuProTyrlGlu 232  
 |||||  
 338 GACCTGTGTTTGACGAGACCTTCACTTCTATGTCATCCCTACACCA 367  
 232 nIleGlnGluLeuAlaLeuHisPheThrIleLeuSerPheAspArgPheS 249  
 |||||  
 368 GCTCAGAGACTGTGTGCTGCTCTTCTCTCAGCTTTGACCGCTTCT 437  
 249 eArgAspAspIleIleGlyGluValLeuIleProLeuSerGlyIleGlu 265  
 |||||  
 438 CTCGGAGATGATGATGGAGGTGATGCTCCTGCTGACGAGGTGGAC 487  
 266 LeuSerGluGlyLysMetLeuMetAsnArgGluIleIleLysArgAsnVa 282  
 |||||  
 488 CCCAGCACAGGCAAGGTACAACTGACACAGGAC.ATCATCAAAAGAAATAT 536  
 282 lArgLysSerSerGlyArgGlyGluLeuLeuIleSerLeuGlySerGlns 299  
 |||||  
 537 CCAAAATGCATCAGCAGAGGAGGAGCTCCAGCTGCTCTGCATATATAC 586  
 299 eThrThrAsnThrLeuThrValValLeuLysAlaArgHisLeuPro 315  
 |||||  
 587 CTGGGACACAGAGAAATGACAGTGTGCTCTCAAGCCAGACACTTCCCG 636  
 316 LysSerAspValSerGlyLeuSer...AspProTyrlValLysValAsnLe 331  
 |||||  
 637 AAGATGATATACACCGCTCTCAGATATCTTATGCAAGTGAAAGCT 686  
 331 uTyThrAlaLysLysArgIleSerLysLysThrHisValLysLysC 348  
 |||||  
 687 CTACTAGCGCAGAAAGCGATTGC.AAGAGAAAAACCATGTGCAAGAGT 735  
 348 yThrProAsnAlaValAlaPheAsnGluLeuPhe 358  
 |||||  
 736 GCACCTTTGAAGCCCATCTTCAATGAATCTTT 767

seq\_name: gb\_est2:BF937960

```

seq_documentation_block:
LOCUS      BF937960      671 bp      mRNA      22-JAN-2001
DEFINITION fml0905.v1 zebrafish adult retina cDNA Danio rerio cDNA clone
            4200777.5 similar to TR:Q9RON3 Q9RON3 SYNAPTOTAGMIN XI. ;, mRNA
            sequence.
ACCESSION  BF937960
VERSION    BF937960.1 GI:12355280
KEYWORDS   EST.
SOURCE     zebrafish.
            Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Rasbora; Danio.
REFERENCE  Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
            S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood
            K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Persson, B.,
            Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Rittner, E.,
            Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
            and Wilson, R.
            Washu zebrafish EST Project 1998
TITLE      Unpublished (1998)
JOURNAL    Contact: Stephen L. Johnson
COMMENT    Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            Library constructed by: Susan E. Brockerhoff DNA Sequencing by:
            Washington University Genome Sequencing Center Clone distribution:
            ResourceCenter@umprimar Datenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Seq primer: '3' ER from Amersham
            High quality sequence stop: 420.
FEATURES   Location/Qualifiers
            source          1..671
                        /organism="Danio rerio"
                        /strain="wild-type"
                        /db_xref="taxon:7955"
                        /clone="4200777"
                        /clone_lib="Zebrafish adult retina cDNA"
                        /sex="mixed"
                        /dev_stage="1-2 years"
                        /lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')"
                        /note="Vector: Lambda ZAP II (PBLuescript SK-); Site_1:
                        EcoRI; Site_2: SalI; This zebrafish library was
                        constructed by Dr. Susan E. Brockerhoff (email:
                        sbrocker@u.washington.edu) RZPD library number: 760"
BASE COUNT 171 a 179 c 168 g 152 t 1 others
ORIGIN
alignment_scores:
            Quality: 700.00      Length: 213
            Ratio: 3.723      Gaps: 2
Percent Similarity: 88.263 Percent Identity: 61.502
alignment_block:
US-09-680-121-2 x BF937960 ..
Align seg 1/1 to: BF937960 from: 1 to: 671
154 LeuGlyThrIeuPheSerLeuGlyTyrAsnPhcGluArgLysAlaPh 170
|||||
21 CTGGGGACGCTCAGTTGCTATGTACTACACTCCCAAAAAAGCCCT 70
|||||
170 eValaIAsnIleLysGluAlaArgLysProAlaMetAspGlnGln 187
|||||
71 GGTGCTGACTATCATAGGGCGCGAGGGGCTCCCGACGTGATGATGACAGG 120
|||||
187 erMetThrSerAspProTyrIleLysMetThrIleLeuProGluLys 203
|||
121 CAGGCACTCTGACCCCTATGTGATGATGACCATTTGCTGAGAAAGAG 170

```

```

204 HisLysValIleThrArgValIleuArgLysThrLeuAspProAlaPheAs 220
|||||
171 CACCGGTTAAAGACGGGTCTCTGAGGAGACATCTGAGCGCTGTGTA 220
|||||
220 pGluThrPheThrPheTyrGlyIleProTyrThrGlnIleGlnIleuA 237
|||||
221 CGAAGCTTCATCTTCTACGGCGTCCCATATGACCTGCTGCCAGAGCTTA 270
|||||
237 lAluHisPheThrIleLeuSerPheAspArgPheSerArgAspAspIle 253
|||||
271 CACTGCATCTTCGATGCTCCTCACCCTTGATCGTTTGACCGGATGATGTC 320
|||||
254 lIleGlyIleValIleuIleProLeuSerGlyIleGluLeuSerGluGly 270
|||||
321 ATCGGGAGGACAGTGGTCCGTTAGAGGCGGTGATCCACGACGGGAG 370
|||||
270 smetLeuMetAsnArgGluIleIleLysArgAsnValArgLysSerG 287
|||||
371 AGGCAATCACCACCAATATGACAGAGAGACACACTGTGAAGC. 419
|||||
287 lYArgGlyGluLeuLeuIleSerLeuCysTyrGlnSerThrAsnThr 303
|||||
420 . CGTGGGCAATTGCTGCTCTCTGCTCTATCATCCTGTACTACACAG 467
|||||
304 LeuThrValValIleuLysAlaArgHisLeuProLysSerAspAla 320
|||||
468 CTTAATGTGGTTGCTTAAGGCCAAACACCTTCAACATGACATGCAC 517
|||||
320 rGlyLeuSer...AspProTyrValLysValAsnLeuTyrHisAlaLys 336
|||||
518 TGGCTCTCAGCAGCACCCCTATGTAAAGTAAAGCAGTCTTACGGCGCA 567
|||||
336 yAsrGlyLeuSerLysLysLysThrHisValLysLysCysThrProAsnAla 352
|||||
568 AACGCATACCCCAAGAAAGAAACACACGCTGAAGAAATCACATCAATCCC 617
|||||
353 ValPheAsnGluLeuPheValPheAspIleProCysGlu 365
|||||
618 GTATTCATGATTCATTATTCATGATTCCTGCTGAG 656
|||||
seq_name: gb_est2:BG671553
seq_documentation_block:
LOCUS      BG671553      791 bp      mRNA      30-APR-2001
DEFINITION DRNBUC05 Rat DRG Library Rattus norvegicus cDNA clone DRNBUC05 5',
            mRNA sequence.
ACCESSION  BG671553
VERSION    BG671553.1 GI:13893652
KEYWORDS   EST.
SOURCE     Norway rat.
            Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 791)
            Xiao, H.-S., Han, Z.-G., Zhang, F.-X., Huang, Q.-H., Lu, Y.-J., Bao, L., Fu, G.,
            Guo, C., Yan, Q., Jin, S.-X., Zhu, Z.-D., Xu, X.-R., Li, N.-G., Chen, Z. and
            Zhang, X.
            Distinct gene expression profiles of rat dorsal root ganglion
            induced by peripheral nerve axotomy
            Unpublished (2001)
JOURNAL    Contact: Zhang Xu
COMMENT    Laboratory of Sensory System
            Institute of Neuroscience
            320 Yue Yang Road, Shanghai 200031, P.R.China
            Tel: 86-21-64748700-121
            Fax: 86-21-64713446
            Email: xu.zhang@ion.ac.cn
            This clone is also available at Chinese National Human Genome
            Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park,
            Pudong New Area, P.R.China. Please contact with Zhang Xu
            (xu.zhang@ion.ac.cn) or Han Zeguang (hanzg@hgc.sh.cn)

```



```

322 euSerAspProTyrValIysValAsnLeuTyrHisAlaLysArgIle 338
|||||
325 TTTCAGATCCCTACGTCAAGTAGACTGTACCATGCCAGAGAGATC 276
|||||
339 SerLysLysLysThrHisValIysLysCysThrProAsnAlaValPheas 355
|||||
275 TCTAAAGAAAGAGACTCACGTGAAAAAGTGCATCCCAACCCAGTGTCAA 226
|||||
355 ngLuleuPheValPheAspIleProCysGluGlyLeuGluAspIleSerV 372
|||||
225 CGACGCTTTGCTCTGTGATATCTCTGTGAGAGCTTTGAAGAAATAAGTG 176
|||||
372 alGluPheLeuValIeuAspSerGluArgLysSerArgAsnGluValIle 388
|||||
175 TTGAATTTTAAAGTTTGATTTGATTCGAAAGGGATCCGCAATGAGGTATC 126
|||||
389 GlyGluLeuValLeuGluAlaAlaIaGluGlyThrGlyGlyLysIleHis 405
|||||
125 GGCGGCTGTGCTGGGTGCCACAGCAAGGAAGTGTGGGGGCATCTG 76
|||||
405 pLysGluIleCysAspTyrProArgArgGlnIleAlaLysThrHisValI 422
|||||
75 GAAGGAGATCTGTGACTTCCCAAGACAAATGCTAACTGCATATTC 26
|||||
422 euCysAspGly 425
|||||
25 TCTGTATGTGT 15

```

```

seq_name: gb_est1:AW367632
seq_documentation_block: 589 bp mRNA EST 04-FEB-2000
LOCUS AW367632
DEFINITION MR0-HT0165-121199-003-a09 HT0165 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW367632
VERSION AW367632.1 GI:6872282
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 589)
AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR0&cl2=MR0-HT0165-
121199-003-a09&cl3=1999-11-12&cl4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 563.
Location/Qualifiers
1..589
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0165"
/dev_stage="Adult"
/note="Organ: head,neck; Vector: puc18; Site: 1; Smal;
Site:2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

```

FEATURES
Source

```

```

BASE COUNT 149 a 161 c 154 g 125 t
ORIGIN
alignment_scores:
Quality: 668.00 Length: 184
Ratio: 4.098 Gaps: 1
Percent Similarity: 88.587 Percent Identity: 66.304
alignment_block:
us-09-680-121-2 x AW367632 ..
Align seq 1/1 to: AW367632 from: 1 to: 589

```

```

190 SerAspProTyrIleLysMetThrIleLeuProGluLysHisLysVal 206
|||||
37 TCTGACCCCTAACATCAAAATGACCATCTTCTGCAAAAGGACATCGGGT 86
|||||
206 LysThrArgValLeuArgLysThrLeuAspProAlaPheAspGluThrP 223
|||||
87 GAAGACCAAGAGTCTGCGGAGAACCTTGACCTGTGTTGACGAGACT 136
|||||
223 heThrPheTyrGlyIleProTyrThrGlnIleGlnLuleuAlaLeuHis 239
|||||
137 TCATCTCTTAATGGCATCCCTACAGCCAGCTGCAGAGCTGTGTCGCAC 186
|||||
240 PheThrIleLeuSerPheAspArgPheSerArgAspIleIleGlyI 256
|||||
187 TTCCTGTCTCTACACTTGAACCGCTCTCTCGGATATGATCATTTGGCGA 236
|||||
256 uValLeuIleProLeuSerGlyIleGluLeuSerGluLysMetLeu 273
|||||
237 GGTATGATGTCGACTGCGAGGGGTGGACCCAGCAGCAGGATACAC 286
|||||
273 etAsnArgGluIleIleLysArgAsnValArgLysSerSerGlyArgGly 289
|||||
287 TGACCAAGGAGCATCATCAAAAGCAATATCCAGAAAGTGCATCAGCAGAGG 336
|||||
290 GluLeuLeuIleSerLeuCysTyrGlnSerThrThrAsnThrLeuIle 306
|||||
337 GAGTCCAGGCTGTCTGTCTATATACACCTGTGACAGAGAAAGACAGT 386
|||||
306 lValValLeuLysAlaArgHisLeuProLysSerArgValSerGlyLeu 323
|||||
387 GGTGTCTTCAAGCCAGACACTTGGCCGAGAGATATACACGGTCTCT 436
|||||
323 er.....AspProTyrValIysValAsnLeuTyrHisAlaLys 335
|||||
437 CAGGTAGGGGCCCGCATCTTATGTCAAGTGAGACGTCCTACTACGGCAGA 486
|||||
336 LysArgIleSerLysLysLysThrHisValLysLysCysThrProAsnAl 352
|||||
487 AAGGCACTTGGCAAGAAAGAACCCATGTGCAGAAAGCATCTTTGGACCC 536
|||||
352 aValPheAsnGluLeuPheValPheAspIleProCysGluGlyLeuGlu 369
|||||
537 CATCTTAATGATGCTTTCATTAAGACAGTCCAGTGCATGCTGTGCTG 586
|||||
369 sp 369
||
587 AT 588

```

```

seq_name: gb_est1:AW367662
seq_documentation_block: 589 bp mRNA EST 04-FEB-2000
LOCUS AW367662
DEFINITION MR0-HT0165-191199-003-a09 HT0165 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW367662
VERSION AW367662.1 GI:6872312
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



```

294 SerLeuCystrGlnSerThrThrAsnThrLeuThrValValValleu 310
    |||
78  TCCTCTGTTATCAGTCACCTACCAACAGCCTCCTGCTGCTTAA 127
    |||
310 salAArgHisLeuProLysSerAspValSerGlyLeuSerAspProTyr 327
    |||
128 AGCGCGGACCTACCGAAATCTCATGTGCTTGAGATCCCTACG 177
    |||
327 allYsValAsnLeuTyrHisAlaLysLysArgLleSerLysLysThr 343
    |||
178 TCAAAGTGAACCTGTACCATGCCAGAGAGAAATCTCTAAAAAGAGACT 227
    |||
344 HisValLysLysCysThrProAsnAlaValPheAsnGluLeuPheValPh 360
    |||
228 CAGCTGAAAAAGTGCACCTCCAGCAGAGGTGTCAAGCAACTGTTGCTT 277
    |||
360 eaSPILeProCysGluGlyLeuGluAspLleSerValGluPheLeuVal 377
    |||
278 TGATATTCCTGTGAGAGCTTGAGAAATAGTGTGAATTTTGTGTT 327
    |||
377 euAspSerGluArgGlySerArgAsnGluValLleGlyGlnLeuValLeu 393
    |||
328 TGGAATCTGAAAGGGATCCGAAATGAGGTGATCGGCGGTTGCTCTG 377
    |||
394 G1YAlAlAlAlaGluGlyThrGlyGlyLuhHisTrpLysGluLleCysAs 410
    |||
378 GGTCGCCACAGCAGAGAAAGTGTGGGGGCGCACTGCAGAGAGATCTGTGA 427
    |||
410 pTYrProArgArgGlnLleAlaLysTrpHisValLeuCysAspGly 425
    |||
428 CTGCCGAGGAGACAAATGCTAAGTGCATATGCTGTGATGTGT 473
    |||
seq_name: gp_est1:AW914163

```

```

seq_documentation_block:
LOCUS      AW914163          712 bp      mRNA      EST      25-MAY-2000
DEFINITION EST345467 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
            RGIAB26 5' end, mRNA sequence.
ACCESSION  AW914163
VERSION    AW914163.1 GI:8079837
KEYWORDS   EST.
SOURCE     Rattus sp.
ORGANISM   Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 712)
AUTHORS   Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
            Kerlavage,A.R. and Adams,M.D.
TITLE      Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
            Gene Index
JOURNAL    Unpublished (1998)
COMMENT    Other_ESTs: EST345466
            Contact: Lee, NH
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@igf.org
            This clone is available through the ATCC, contact the ATCC
            tel#703-365-2700 for further information
            Seq primer: M13 Reverse.
FEATURES
SOURCE     Location/Qualifiers
            1..712
             /organism="Rattus sp."
             /db_xref="taxon:10118"
             /clone="RGIAB26"
             /clone_lib="Normalized rat brain, Bento Soares"
             /note="Organ: brain; Vector: pT7T3pac; Site_1: EcorI;
             Site_2: NotI"

```

```

BASE COUNT      185 a      145 c      178 g      204 t
ORIGIN

```

```

alignment_scores:
    Quality: 665.00      Length: 132
    Ratio: 5.076
    Percent Similarity: 99.242      Percent Identity: 93.939
alignment_block:
US-09-680-121-2 x AW914163 ..
Align seg 1/1 to: AW914163 from: 1 to: 712

```

```

294 SerLeuCystrGlnSerThrThrAsnThrLeuThrValValValleu 310
    |||
78  TCCTCTGTTATCAGTCACCTACCAACAGCCTCCTGCTGCTTAA 127
    |||
310 salAArgHisLeuProLysSerAspValSerGlyLeuSerAspProTyr 327
    |||
128 AGCGCGGACCTACCGAAATCTCATGTGCTTGAGATCCCTACG 177
    |||
327 allYsValAsnLeuTyrHisAlaLysLysArgLleSerLysLysThr 343
    |||
178 TCAAAGTGAACCTGTACCATGCCAGAGAGAAATCTCTAAAAAGAGACT 227
    |||
344 HisValLysLysCysThrProAsnAlaValPheAsnGluLeuPheValPh 360
    |||
228 CAGCTGAAAAAGTGCACCTCCAGCAGAGGTGTCAAGCAACTGTTGCTT 277
    |||
360 eaSPILeProCysGluGlyLeuGluAspLleSerValGluPheLeuVal 377
    |||
278 TGATATTCCTGTGAGAGCTTGAGAAATAGTGTGAATTTTGTGTT 327
    |||
377 euAspSerGluArgGlySerArgAsnGluValLleGlyGlnLeuValLeu 393
    |||
328 TGGAATCTGAAAGGGATCCGAAATGAGGTGATCGGCGGTTGCTCTG 377
    |||
394 G1YAlAlAlAlaGluGlyThrGlyGlyLuhHisTrpLysGluLleCysAs 410
    |||
378 GGTCGCCACAGCAGAGAAAGTGTGGGGGCGCACTGCAGAGAGATCTGTGA 427
    |||
410 pTYrProArgArgGlnLleAlaLysTrpHisValLeuCysAspGly 425
    |||
428 CTGCCGAGGAGACAAATGCTAAGTGCATATGCTGTGATGTGT 473
    |||
seq_name: gp_est1:AU035592

```

```

seq_documentation_block:
LOCUS      AU035592          813 bp      mRNA      EST      12-JUL-2000
DEFINITION AU035592 Sugano mouse brain mncb Mus musculus cDNA clone MNCB-0621,
            mRNA sequence.
ACCESSION  AU035592
VERSION    AU035592.1 GI:3718595
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 813)
AUTHORS   Sasaki,Z., Suzuki,Y., Watanabe,M., Inai,J., Shibu,A., Yoshida,K.,
            Hata,H., Yamaguchi,R., Tateyama,S. and Sugano,S.
TITLE      Construction of mouse full length-enriched cDNA libraries
JOURNAL    Unpublished (1998)
COMMENT    Contact: Katsuyuki Hashimoto
            Division of Genetic Resources
            National Institute of Infectious Diseases
            23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            Email: khash@nih.go.jp
            POLYA=No.
FEATURES
SOURCE     Location/Qualifiers
            1..813
             /organism="Mus musculus"
             /strain="C57BL"
             /db_xref="taxon:10090"

```





---

•  
•  
•  
•



OM of: US-09-680-121-2 to: Issued\_patents\_NA.\* out\_format: pfs  
Date: Dec 13, 2001 12:00 AM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODEL=frim+p2n.model -DEV=xip  
-Q/cgn2\_1/USPTO.spool/US09680121/runat\_12122001\_105512\_22747/app\_query.fasta.1.488  
-DB=Issued\_Patents\_NA -OPMT=fastp -SUFFIX=p2n.rn1  
-GAPOP=12.000 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000  
-LOOPEXT=0.000 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000  
-XGAPEXT=0.500 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000  
-YGAPEXT=0.500 -DELOP=6.000 -DELEXT=7.000 -STARF=1  
-MATRIX=blotsum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200  
-THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US09680121.ecgn1\_1\_81 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
-WAIT -THREADS=1

## Search information block:

Query: US-09-680-121-2  
Query length: 425  
Database: Issued\_Patents\_NA.\*  
Database sequences: 351203  
Search time (sec): 88.380000

## score list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-036-315-1		2201.00	4677.32	7.0e-253	3891
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-872-979-4		338.50	690.79	7.9e-31	1207
/cgn2_6/ptodata/2/ina/6A.COMB.seq:3266464-1		165.50	309.33	1.4e-09	2599
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-094-714A-48		161.00	301.53	3.8e-09	2244
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-225-749-24		161.00	301.53	3.8e-09	2245
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-609-049A-29		158.00	284.16	3.5e-08	5285
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-170-996-29		158.00	284.16	3.5e-08	5285
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-609-049A-27		157.00	278.74	7.1e-08	6831
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-170-996-27		157.00	278.74	7.1e-08	6831
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-539-205A-5		135.00	244.58	5.7e-06	3148
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-465-961A-16		100.50	154.27	0.6075	8700
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-909-954-3		132.00	235.04	1.9e-05	3117
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-895-801-2		117.00	204.22	0.0010	2750
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-642-846-1		102.50	165.14	0.1506	5154
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-834-625-16		100.50	154.27	0.6075	8700
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-465-961A-16		100.50	154.27	0.6075	8700
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-663-998-2		98.50	154.81	0.5665	5952
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-804-196-1		97.50	150.76	0.9523	6909
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-658-340-1		97.50	150.76	0.9523	6909
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-746-111-26		97.50	150.76	0.9523	6909
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-035-360-1		97.00	154.19	0.6139	4856
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-035-360-1		97.00	153.88	0.6380	4972
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-001-951-17		97.00	153.88	0.6380	4972
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-4		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-5		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-4		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-5		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-4		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-5		96.00	166.87	0.1206	1518
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-1		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-3		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-1		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-3		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-3		96.00	165.03	0.1527	1753
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-11		96.00	164.91	0.1551	1770
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-13		96.00	164.91	0.1551	1770
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-11		96.00	164.91	0.1551	1770
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-13		96.00	164.91	0.1551	1770

/cgn2\_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-11 + 96.00 164.91 0.1551 1770  
/cgn2\_6/ptodata/2/ina/6B.COMB.seq:US-09-397-979-13 - 96.00 164.91 0.1551 1770  
/cgn2\_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-25 + 96.00 162.90 0.2009 2073  
/cgn2\_6/ptodata/2/ina/6B.COMB.seq:US-08-929-501-26 - 96.00 162.90 0.2009 2073  
/cgn2\_6/ptodata/2/ina/6A.COMB.seq:US-09-140-177-25 + 96.00 162.90 0.2009 2073

## seq\_name: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:US-09-036-315-1

## seq documentation block:

Sequence 1, Application US/09036315  
Patent No. 6218523  
GENERAL INFORMATION:  
APPLICANT: French, Cynthia K.  
APPLICANT: Schneider, Patrick A.  
APPLICANT: Yamamoto, Karen K.  
TITLE OF INVENTION: Prostate Cancer-Specific Marker  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/036,315  
FILING DATE: 06-MAR-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/047,811  
FILING DATE: 15-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/041,246  
FILING DATE: 07-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Scirella, John R.  
REGISTRATION NUMBER: 32,944  
REFERENCE/DOCKET NUMBER: 018002-000210US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3891 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 151..1425  
OTHER INFORMATION: /product= "Repro-PC-1.0"  
US-09-036-315-1  
alignment\_scores:  
Quality: 2201.00 Length: 425  
Ratio: 5.179 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-680-121-2 x US-09-036-315-1 ..  
Align seg 1/1 to: US-09-036-315-1 from: 1 to: 3891  
1 MetAlaProlethrThrSerArgGluGluPheAspGluIleProThrVa 17  
|||||  
151 ATGCTCCGATCACACACCGCGGAGAAATTGATGAAATCCCCACAGT 200

```

17  lvalgllylepheseralephleuvalphehrvalserleuhea 34
201  GGTGGGATCTTCAGTCATTTGGCTGCTTCACACTCTCTCTCTTGG 250
34  latrpillecyscgluarlyserSerLysSerAsnLysThrProPro 500
251  CATGATCTGCTGCAGAGAAATCATCCAACTCAACAAACCTCCCA 300
51  TyrLysPheValHisValLeuLysGlyValAspLetyrProGluAsnLe 67
301  TCAAGATTGTGTCATGTGCTTAAGGGAGTGTGATATTTACCTGAAACCT 350
67  uasnserLysLysPheGlyAlaAspAspLysAsnGluValLysAsnL 84
351  AATAGCAAAAAGAGTTTGAGCAGCATGATAAAATGAAGTAAGATTA 400
84  yspProAlaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAsp 100
401  AGCCACCTGTCCTCCAAAGATTCATTCATGATGATGATAAAGAGAT 450
101  LeuAsnGlyAsnPheProLysThrAsnLysProGlySerProSerAs 117
451  CTCAAATGGCAATTTCCAAACCAACCAACCTGAGCTGCTCTG 500
117  PleuGluAsnAlaThrProLysLeuPheLeuGluGlyGlyLysGlySerV 134
501  TCTGGGGAATGCACCCGGAAGCTCTTTTGAAGGGCAAAAGACTCAG 550
134  alserProGluSerLeuLysSerSerThrSerLeuThrSerGluLys 150
551  TTTTCCCTGAGATTAAAGTCACACCTTCCTTACTTCAGAAAGAAA 600
151  GluGluLysLeuGlyThrLeuPheSerLeuGluLysThrAspPheGluThr 167
601  CAGAGAGAGCTGGAGACTCTCTCTCTCTCTAGATACAACTTCAGAG 650
167  gLyAlaPheValValAsnLysGluAlaArgGlyLeuProAlaMetA 184
651  AAAAGCAATTTGTGTCAATATCAGAGAGCCCTGCTTCCAGCCATCG 700
184  spgGluInserMetThrSerAspProTyrLysMetThrLeuPro 200
701  ATGAGCAGTCGATGACCTGCTGACCATATCAAAATGAGATCTCCCA 750
201  GluLysLysHisLysValLysThrArgValLeuArgLysThrLeuAspPr 217
751  GAGAGAGAGCATTAAGTGAACCTAGAGTGTGAGAAAAACCTTGATCC 800
217  calPheAsnGluThrPheThrPheTyrGlyLeuProTyrThrGluLeu 234
801  AGCTTTTGATGACCTTTCATTCATTCATGGGATACCCACCAATTC 850
234  lngLysAlaAlaLeuHisPheThrLeuSerPheAspArgPheSerArg 250
851  AAGAAATGGCTTGCACTTCACAAATTTGAGTTTGAAGGTTTTCAGA 900
251  AspAspLysLeuGlyValLeuLeuLeuProLysSerGlyLeuLeuSe 267
901  GATGATATCATTTGGGAGTTTAATTCCTCTCGGGAATTTGAATATTC 950
267  rGluGlyLysMetLeuMetAsnArgLysLeuLysArgAsnValArgL 284
951  TGAAGCAAAAATGTTATGAAATGAGATCATCAAGAAATGTTAGGA 1000
284  ySerSerGlyArgGlyLysLeuLeuLeuSerLeuCysTyrGlySerThr 300
1001  ACTCTTCAGAGAGGGGTGAGTACTGATCTCTCTGATACAGCCACC 1050
301  ThrAsnThrLeuThrValValLeuLysAlaArgHisLeuProLysSe 317
1051  ACAAACACTTAAGTGTGTTGCTTAAAGCTCGACATCTGCTTAATTC 1100
317  rAspValSerGlyLeuSerAspProTyrValLysValAsnLeuThrHisA 334
1101  TGAATGTCCGGACTTTCAAGATCCATATGTCAAAGTCAACCTGATCAG 1150
334  lalysLysArgLysSerLysLysThrHisValLysLysCysThrPro 350
1151  CCAAAAGAGAAATCTCCAAAGAGACATCATGTGAAGAAATGCACCC 1200
351  AsnAlaValPheAsnGlyLeuPheValPheAspLysProCysGluGlyLe 367
1201  AATGAGATTTCATATGAGCTGTTGCTTGTGATATTCCTGTGAGGGCT 1250
367  uGluAspLysSerValGluPheLeuValLeuAspSerGluArgLysSerA 384
1251  TGAAGATTAAGTGTGAAATTTTGGTTTGGATTCTGAAAGGGGCTCC 1300
384  rAsnGluValLysGlyLysLeuValLeuGlyAlaAlaLysGluLysThr 400
1301  GAAATGAGGTAAATCGGGAGTGTAGTCTTGGTGCACACCAAGAGAACT 1350
401  gLyGlyGluHisThrPlyGlyLysCysAspTyrProArgArgGluLeuAl 417
1351  GTGGAGACACTGGAAAGAGATCTGTGACTACCCGAGAGCAAAATTC 1400
417  alYsrPHisValLeuCysAspGly 425
1401  CAAGTGGCAGCTGCTCTGTGATGCT 1425

seq_name: /cgn2_6/plodata/2/lna/6A.COMB.seq:us-08-872-979-4
seq_documentation_block:
: Sequence 4, Application us/08872979
: Patent No. 6074844
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Lal, Preeti
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: TWO NEW HUMAN MEMBRANE FUSION PROTEINS
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS windows Version 2.0
: SOFTWARE: FASTSEQ FOR windows
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/872,979
: FILING DATE: December 5/4
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REFERENCE/DOCKET NUMBER: 36,749
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELETYPE:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1207 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:

```

```

584 CCAAAAAGGCGACGGCGCCCTACTTCAATAGAGCCCTTCACTTCTG 903
362 ILpOCysgInglYleuGlnaAspILeSerValGluPheLeuValLeuAs 378
904 GTCCCTTCAGCCAGGAGCCAGAAATGTGACTGTGCTGCTGCTGTGGA 953
378 pSerGluArgGlySerArgasnGluValILlegIghInLeuValLeuGLA 395
954 CCGACACCTGCGCTCCGAGCTAGAGCCCGTAGGCAAGGTGCACCTGGTG 1003
1004 CCGGCGCCTCGGGGCGACGCCCTCGAGCAGCTGGGAGACATGCTGGCCAC 1053
412 ProArgArgInIle 416
1054 GCAGCCAGGAGGAGTG 1068

seq_name: /cgn2.6/plodata/2/ina/backfiles1.seq:5266464-1

seq_documentation_block:
Patent No. 5266464
APPLICANT: HOUSEY, GERARD
TITLE OF INVENTION: METHOD OF SCREENING FOR PROTEIN INHIBITORS
AND ACTIVATORS
NUMBER OF SEQUENCES: 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/392,073
FILING DATE: 10-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 154,206
FILING DATE: 10-FEB-1988
;SEQ ID NO:1
;LENGTH: 2599
5266464-1

alignment_scores:
Quality: 165.50 Length: 157
Ratio: 1.655 Gaps: 7
Percent Similarity: 63.694 Percent Identity: 28.025

alignment_block:
US-09-680-121-2 x 5266464-1 ..

Align seg 1/1 to: 5266464-1 from: 1 to: 2599

120 AsnAlaThrProIysLeuPheLeuGluGlyIysGluSerValSerP 136
|||||
482 AARGCGACCTCTGATGATG.....AATGTCACAA 513
136 ogUserLeuYserSerThrSerLeuThrSerGluGluYsgInglu 153
514 GCGCTGCTATGACGTCCCGACCTCTGTGGCAGCCAGCACACAGAAC 563
153 yslengIythrLeuPheSerLeuGluIyrasnPheGluArgLysAla 169
564 GCGGTGCGCCATCTAC.....ATCAGGCCCACTGACAGGAGAGTTC 607
170 PheValValasnIleIysGluAlaArgGlyLeuProAlaMetaspGlu 186
608 CTCATCTGTTGTAAAGATGCTAAAAATCTGCTACCTAGACGCCCAA 657
186 nserMetThrSerapProIyrIleIysMetThrILleuProGluYsL 203
658 CGGCTT...TCAGATCCCTACGTAAACAGAAACTGATCCCTGTATCCA 704
203 ys.....HisLysValIysThrArgValLeuArgGlyThrLeuAspPro 217
705 AAGGTGAGACAGCAAGACCAAGCAAGACTATCAATATGCTCCTCAACCG 754
218 AlaPheaspGluThrPheThrPheIyrGlyIleProIyrThrGlnIleGI 234
|||||

```

Fri Dec 14 10:32:01 2001

us-09-680-121-2.p2n.rni

Page 4

```
755 GAGTGAACGAACCTTCAGATT.....CAGCTGAA 786
234 nglu.....LeuAlaLeuHisPheThrIleLeuSerPheAspA 247
787 GGAATATAGACAAAGACAGACATCTGCTGAGATCTGGATGTGGACC 836
247 rphSerArgAspAspIleIleGlyGluValLeuIleProLeuSerGly 263
837 TGACCGGACGAGATGATCTGATGAGATCTGCTGCTTGGATTCA... 883
264 IleGluLeuSerGluGlyLys 270
884 ...GAACATGACGAAAGCCGG 901
```

seq\_name: /cgn2\_6/ptodata/2/lna/6A\_COMB.seq:US-09-094-714A-48

```
seq_documentation_block:
; Sequence 48, Application US/09094714A
; Patent No. 6117847
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Nicholas M. Dean
; TITLE OF INVENTION: OLIGONUCLEOTIDES FOR ENHANCED MODULATION OF
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESS: Woodcock Washburn Kurtz Mackiewicz & No. 6117847ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2 PC-DOS
; OPERATING SYSTEM:
; SOFTWARE: WORDPERFECT 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,714A
; FILING DATE: June 15, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/601,269
; FILING DATE: 14-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/478,178
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/089,996
; FILING DATE: 09-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/852,852
; FILING DATE: 16-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2943
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-094-714A-48
```

alignment\_scores:
Quality: 161.00 Length: 177
Ratio: 1.464 Gaps: 7
Percent Similarity: 62.147 Percent Identity: 28.249

alignment\_block:
US-09-680-121-2 x US-09-094-714A-48 ..
Align seg 1/1 to: US-09-094-714A-48 from: 1 to: 2244

```
113 GlySerProSerAspLeuGluAsnAlaThrProLysLeuPhe..... 126
355 GGAAGACCCACCTTCTGCATCAGTGGTGTCTACTGCTCTATGAGACTTAT 404
127 LeuGluGlyGlyLys.....GluSerValSerProGluS 138
405 CCATCAAGGATGAATGATGACACCTGCATATGACGTTTCAACAACT 454
138 erLeuYsserSerThrSerLeuThrSerGluGlyGlnGlyLysLeu 154
455 GCGTCATCATATGCCCCAGCCCTGCGGAATGATGCACACTGAGAGAG 504
155 GlyThrLeuPhePheSerLeuGluLysAsnPheGluArgLysAlaPhe 171
505 GCGCGCATTTACTTAAAGCTAGAGTGTGCTGATGAAG.....CTCA 548
171 lValAsnIleLysGluAlaArgGlyLeuProAlaMetAspGluInSer 188
549 TGTCACAGTACGAGATGCAAAATCTATATCCCTATGATGCCAAAGGGC 598
188 eThrSerAspProTyrIleLysMetThrIleLeuProGluLysLysHis 204
599 TT...TCGATCCTTATGTGAAGCTGAATATTCCTGATCCCAAGAT 645
205 ....LysValLysThrArgValLeuArgLysThrLeuAspProAla 219
646 GAAAGCAAGCAAAACCAAAACCAATCCGTCACACTAATATCCGAGAG 695
219 eAspGluThrPheThrPheThrGlyLysLeuProTyrThrGlnIleGln 236
696 GAATGAGTCCCTTACATTCATTAACCTTCAGACAAAGACCGAGCAG 745
236 euAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArgAsp 252
746 TGCTGTAGAA.....ATCTGGAGCTGGAGTGCACACAGAGAAAGAT 789
253 IleIleGlyGluValLeuIleProLeuSerGlyLys.....Glu 266
790 TTCATGGATGCCCTTCTGAGATTGAGAGTGTGAGATGAGATCCGCGC 839
266 uSerGluGlyLysMetLeuMetAsnArgLys 276
840 CAGTGAATGCTACAAAGTTGCTTACCAAGAA 870
```

seq\_name: /cgn2\_6/ptodata/2/lna/6B\_COMB.seq:US-09-225-749-24

```
seq_documentation_block:
; Sequence 24 Application US/09225749
; Patent No. 6308320
; GENERAL INFORMATION:
; APPLICANT: Dean, Nicholas M.
; TITLE OF INVENTION: Modulation of c-jun using inhibitors of protein kinase C
; FILE REFERENCE: ISIS3313
; CURRENT APPLICATION NUMBER: US/09/225,749
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2245
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (28)..(2046)
; PUBLICATION INFORMATION:
; JOURNAL: Nucleic Acids Res.
; VOLUME: 18
```

```

; ISSUE: 8
; PAGES: 2183
; DATE: 1990-04-25
; DATABASE ACCESSION NUMBER: X52479/genbank
; DATABASE ENTRY DATE: 1993-09-12
; US-09-225-749-24

```

## alignment\_scores:

```

Quality: 161.00      Length: 177
Ratio: 1.464         Gaps: 7
Percent Similarity: 62.147   Percent Identity: 28.249

```

## alignment\_block:

```
US-09-680-121-2 x US-09-225-749-24 ..
```

```
Align seg 1/1 to: US-09-225-749-24 from: 1 to: 2245
```

```

113 GlySerProSerAspLeuGluAsnAlaThrProLysLeuPhe..... 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
355 GGAAGCCGCCACCTTCGATCACTGGGTGATGCTGCTATGAGACTTAT 404
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
127 .LeuGluGlyGluLys.....GluSerValSerProGluS 138
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
405 CCATCAAGGGATGAATGTGACACCTGCGATATGACGTTGACAAAGCAT 454
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 erLeuLysSerSerThrSerLeuThrSerGluGluLysGluGluLysLeu 154
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
455 GCCTCATCAATGTCCCGACGCTCTGCGGATGATGATCAGTCAAGAGAG 504
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155 GlyThrLeuPhePheSerLeuGluTyrAsnPheGluArgLysAlaPheVa 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 GGGGGGATTTAACTAAAGGCTGAGTGTGTGATGAAG.....CTCCA 548
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
171 lValAsnIleLysGluAlaArgGlyLeuProAlaMetAspGluInserm 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
549 TGTCACAGTGCAGATGCAAAATTAATCCCTATGATGCCAAACGGGC 598
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
188 eThrSerAspProTyrIleLysMetThrIleLeuProGluLysLysHis 204
    ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
599 TT...TCAGATCCTTATGTGAGAGCTGAACCTTATCTGATCCCAAGAA 645
    ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205 .....LysValLysThrArgValLeuArgLysThrLeuAspProAlaPh 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
646 GAAAGCAAGCAAAACCAACCATCCGCTCCACATAATCCGCAAGTG 695
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 eAspGluThrPheThrPheTyrGlyIleProTyrThrGluIleGluLul 236
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
696 GAATGAGTCTCTTACATTCAATTGAACCTTCAGACAAAGACGACGAC 745
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236 euAlaLeuHisPheThrIleLeuSerPheAspArgPheSerArgAsp 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
746 TGTCTGTAGAA.....ATCTGGAGCTGGATTCGACACAAAGAAAGAC 789
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253 IleIleGlyLysValLeuIleProLeuSerGlyIle.....GluLe 266
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
790 TTCATGGGATCCCTTCCCTTGGAGTTTCGAGAGTGATGAAGATCCGGC 839
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 uSerGluGlyLysMetLeuMetAsnArgGlu 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
840 CAGTGCATGTACAAATTGCTTAACCAAGAA 870
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
seq_name: /cgn2_6/ptodata/2/lna/5b_COMB.seq:US-08-609-049A-29
```

## seq\_documentation\_block:

```

; Sequence 29, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Molz, Lisa
; APPLICANT: Chen, Yen-Wen
; TITLE OF INVENTION: NO. 5948664e1 PI 3-Kinase Polypeptides
; NUMBER OF SEQUENCES: 32

```

## CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

```

## COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,049A
; FILING DATE: 29-FEB-1996

```

## CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

```
NAME: Dow, Karen B.
```

```
REGISTRATION NUMBER: 29,684
```

```
REFERENCE/DOCKET NUMBER: 2307K-063700US
```

```
TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: 415-326-2400
```

```
TELEFAX: 415-326-2422
```

```
INFORMATION FOR SEQ ID NO: 29:
```

```
SEQUENCE CHARACTERISTICS:
```

```
LENGTH: 5285 base pairs
```

```
TYPE: nucleic acid
```

```
STRANDEDNESS: single
```

```
TOPOLOGY: linear
```

```
MOLECULE TYPE: DNA (genomic)
```

```
FEATURE:
```

```
NAME/KEY: CDS
```

```
LOCATION: 3..5180
```

```
FEATURE:
```

```
NAME/KEY: CDS
```

```
LOCATION: 5183..5195
```

```
FEATURE:
```

```
NAME/KEY: CDS
```

```
LOCATION: 5198..5285
```

```
US-08-609-049A-29
```

## alignment\_scores:

```

Quality: 158.00      Length: 346
Ratio: 0.919         Gaps: 17
Percent Similarity: 49.711   Percent Identity: 23.410

```

## alignment\_block:

```
US-09-680-121-2 x US-08-609-049A-29 ..
```

```
Align seg 1/1 to: US-08-609-049A-29 from: 1 to: 5285
```

```

115 ProSerAspLeuGluAsnAlaThrProLysLeuPheLeuGluGly 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4137 CCTTCAGAGATTGCCAGAACTCAAGATTCAGAGATCGAAATGTGTAG 4186
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
131 sGluSerValSerProGluSerLeuLysSer...SerThrSerLeuHis 147
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4187 AGATGCACTTAGGCCCAACTACAGATGCTGAAGCTACTATTCTTTA 4236
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 erGluGluLysGluGluLysLeuGlyThrLeuPhePheSerLeuGluTyr 163
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4237 CTAGCGTATGAGTCAAGTTTGGGAAGCATT.....GCCAACAAGTTT 4280
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4281 AATTTC.....TTCATTCATAAAGCTTCTCAGCTAGCTTTTTC 4318
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
179 .GlyLeuProAlaMetAspGlu.....G 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
4319 TGGCTTCCTTCCTTATGATGAGCCCATCTTCATTCACCGAAGAAAT 4368
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
186 lnsMetThrSerAspProTyrIleLysMetThrIleLeu..... 199
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

4369 ACTCTTTAGACAGATGGCCGATGCAGGAAGACTCTGTCTTTCACATAT 4418  
||||| :||| | ||||| :||| :|||  
ProctylusLeuHisLysValLysThrArgValle 211  
4419 CATAGAATAATCAACCAGATAAACATATATTATGTGTGGATTCCT 4468  
||||| :||| :||| :||| :|||  
ThirPhetPheTyrgLys 221  
4469 AAGAGAGACACCTTGAAACCATCATTTGATTCGGCAGCATTTGATGAT 4518  
||||| :||| :||| :||| :|||  
ThirPhetPheTyrgLys 228  
4519 TTCAGAACTTCACAATAAGCTCAGTATTAATTTTCTCTTGGAAATTA 4568  
||||| :||| :||| :||| :|||  
TyrThrAlaIleGlnIle 236  
4569 CCTGGCTTCTAATAGATGGTCTTGGAAAGAACACATATAAGATGT 4618  
||||| :||| :||| :||| :|||  
LeuHisPheThrLileLeuSerPheAsp 247  
4619 TGCAGCCAGAGAAAATTGAATTAAACAGTTATTTACAGAGTTTGCTG 4668  
||||| :||| :||| :||| :|||  
AspIleIleGlyValIle 259  
4669 ATCATCATCACAGATGTAGCAGAGTGTGATCTGTGTACTTTTTTCCAC 4718  
||||| :||| :||| :||| :|||  
ProLeuSerGlyIleGluLeuSerGlyLysMetLeuMetAsnArg 276  
4719 CTTTACTCGTATGATGAGAAAGCTGACAGA..... 4748  
||||| :||| :||| :||| :|||  
uIleIleLysArgAsnValArgLysSerSerGlyArg..... 288  
4749 .....ATAGTAGGTCTGACAGTGCACATTCCTCCCTTCAGCC 4782  
||||| :||| :||| :||| :|||  
GlyLysLeuLysLileSerLysCysTygLn 298  
4783 CAACCTGGGCCAAATAGAGAGGACGAGTAATTTCTGTCTTTACCGA 4832  
||||| :||| :||| :||| :|||  
SerThrThrAsnThrIleuThrValValLysLysAlaArgHisLeuPr 315  
4833 AAT.....GGACCTCTTCATCATGGTGATGACATCAAGAATCTTGT 4876  
||||| :||| :||| :||| :|||  
OlySerSerAspValSerGlyLeuSerSerProGlyValLysValAsnLeut 332  
4877 GACTGAAGATGGGGCTGAC...CCAATCCCTATGTCAAAACATACCTGCC 4923  
||||| :||| :||| :||| :|||  
yTrHisAlaLysLysArgLileSerLysLysLysThrHisValLysLys 348  
4924 TTCCAGATACCCCAAAGCTCAAAACGTAACCAAAATTTCACGTAA 4973  
||||| :||| :||| :||| :|||  
ThrProAsnAlaValPheAsnGluLeuPheValPheAsp...IleProcy 364  
4974 ACTAGGAACCCCAATTCATTAAGATGCTGTATATAGGCGATACAGCA 5023  
||||| :||| :||| :||| :|||  
SGLugLYLeuGlnAspLileSerValGlnPheLeuValLeuAspSerGlu 381  
5024 AGAAACTCTGAGCGAGAGAGACTTCACATGATGTACGTACGTACGAGAT 5073  
||||| :||| :||| :||| :|||  
rGlySerArgAsnGluValIleGlyGlnLeuValLeu 393  
5074 CACTGCGGAGAAATTTCTTCTGGGTGGATTAACCTG 5111  
||||| :||| :||| :||| :|||

seq\_name: /cgn2\_6/prodata/2/lna/6B\_COMB.seq.us-09-170-996-29

seq\_documentation block:  
Sequence 29, Application us/09170996  
Patent No 6291220  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Moiz, Lisa  
APPLICANT: Chen, Yen wen  
TITLE OF INVENTION: NO. 6291220el PI 3-Kinase Polypeptides

```

NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Phred/In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,996
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/609,049
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-063700US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2420
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 5285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3..5180
FEATURE:
NAME/KEY: CDS
LOCATION: 5183..5195
FEATURE:
NAME/KEY: CDS
LOCATION: 5198..5285
US-09-170-996-29

alignment_scores:
Quality: 158.00 Length: 346
Ratio: 0.919 Gaps: 17
Percent Similarity: 49.711 Percent Identity: 23.410

alignment_block:
US-09-680-121-2 x US-09-170-996-29 ..

Align seg 1/1 to: US-09-170-996-29 from: 1 to: 5285
115 ProSeGaApLeWGuASnAlaThrProLysLeuPheLeuGlulylGly 131
||||| ||| :::: ||| ||:::|
4137 CCTTAGAGATTGCCAGAACTCACAGATTCACAGACTGCAATATGCTAG 4186
131 sCLeuSerValserProGluSerLeuysSer...SerThrSerLeuThrS 147
:::::|:::::|:::::|:::::|:::::|:::::|:::::|:::::|
4187 ACATGCACTACAGCCCAACATACACAGATCGTCGACGACGACGATTTCTT 4236
147 erGlulGluGlnGlnLulysLeuGlyThrLeuPhePheSerLeuLufyr 163
:::::|:::::|:::::|:::::|:::::|:::::|:::::|:::::|
4237 CTAGGCTGATTGAGCAAGTATTGGAGCAATT.....GCCACAAAGTTT 4280
164 AsnPheGluArgLysAlaPheValAsnLleLysGluAlaArg..... 178
||||| |||::: |||:::| ::::| |||
4281 AATTTC.....TTGATTCAATACCTTGCTGACCTACGTTTTC 4318
179 .GlyLeuProAlaMetAspGlu.....G 186

```



```

4319 TGGCCTTCCTTAATGATGAGCCCATCTTCATCTCACCAGAAACAT 4368
|||||
186 InsertMetHrSerAspProTyrIleLeuSerMetHrIleLeu..... 199
|||||
4369 ACTGCTTTAGACAAGATGCCGGATCAGAGAAAGTCTCTGTTTCACATAT 4418
|||||
200 .....ProGluLysHisLysValLysThrArgValLe 211
|||||
4419 CATAGAAATACAAACCAGATTAACATATATATATGTTGTTGAATTC 4468
|||||
211 uArgLys...ThrLeuAspProAla.....PheAspLys 221
|||||
4469 AAGAGAAGACACTTGAAACCATCATTTGATTCGGACATTTGATGAAT 4518
|||||
222 .....ThrPheTrpMetTyrGlyIle 228
|||||
4519 TTCAGGAATTCACAAATAGCTCAGATATATTTTCTCTTGTGAAATTA 4568
|||||
229 Pro.....TyrThrGlnIleGlnIle 236
|||||
4569 CCTGGCTTCTTAATGATGTTCTTGGAGAAGACACATTAAGATGT 4618
|||||
236 uAla.....LeuHisPheTrpIleLeuSerPheAsp 247
|||||
4619 TGCAGCAAGAGAAATGAATTAACACTTATTTACAGAGTTTGATGA 4668
|||||
247 rPheSerArgAsp.....AspIleIleGlyIleValLeuIle 259
|||||
4669 ATCATCAACAGATGTAGCAGAGTGTGATCTGTGTACTTTTTCAC 4718
|||||
260 ProLeuSerGlyIleGluLeuSerGlyLysMetLeuMetLysArgL 276
|||||
4719 CCTTACTTCTGTGATGAGAAAGCTGAAGA..... 4748
|||||
276 uIleIleLysArgAsnValArgLysSerSerGlyArg..... 288
|||||
4749 .....ATAGCTAGGTCTGCAGGTGCAGTCCCTCAGCC 4782
|||||
289 .....GlyGluLeuLeuIleSerLeuCysTyrGln 298
|||||
4783 CAACCTGGGCCCAATAGAGAGACAGTGAAGTTATCTGTTCTTACGA 4832
|||||
299 SerThrThrAsnThrLeuThrValValLeuLysAlaArgHisLeuPr 315
|||||
4833 AAT.....GGCACTCTTCATCATGATGATGCACATCAAGATCTTGT 4876
|||||
315 oLysSerAspValSerGlyLeuSerAspProTyrValLysValAsnLeuT 332
|||||
4877 GACTGAAGATGGGGCTGAC...CCAATCCCTATGTCAAAACATACCTGC 4923
|||||
332 yHisAlaLysLysArgLysSerLysLysThrHisValLysLysCys 348
|||||
4924 TTCACAGTACCCACCAAAACGTAAACGTAACAAATTCACGTAAA 4973
|||||
349 ThrProAsnAlaValPheAsnGluLeuPheValPheAsp...IleProCys 364
|||||
4974 ACTAGGAAACCCCAACATTCATGAATGCTGTGTATATAGTGATACAGCAA 5023
|||||
364 sGluGlyLeuGluAspLysPheSerValGluPheLeuValLeuAspSerGlu 381
|||||
5024 AGAAACTCTAGGACGAGAGAACTCACTCACTGAGTCTCAGTCCAGAAAT 5073
|||||
381 rGlySerArgAsnGluValIleGlyLysLeuValLeu 393
|||||
5074 CACTGCGGAGAAATTTCTTCTGCGTGGAATAACCCCTG 5111
|||||
seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:us-08-609-049A-27
seq_documentation_block:
; Sequence 27, Application US/08609049A
; Patent No. 5948664
; GENERAL INFORMATION:

```

```

/ APPLICANT: Williams, Lewis T.
/ APPLICANT: Moliz, Lisa
/ APPLICANT: Chen, Yen-Men
/ TITLE OF INVENTION: No. 5948664el PI 3-kinase Polypeptides
/ NUMBER OF SEQUENCES: 32
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, 8th Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/609,049A
/ FILING DATE: 29-FEB-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Dow, Karen B.
/ REGISTRATION NUMBER: 29,684
/ REFERENCE/DOCKET NUMBER: 2307K-063700US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-326-2400
/ TELEFAX: 415-326-2422
/ INFORMATION FOR SEQ ID NO: 27:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 6831 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 148..5775
/ US-08-609-049A-27

alignment_scores:
Quality: 157.00 Length: 279
Ratio: 0.981 Gaps: 10
Percent Similarity: 57.348 Percent Identity: 22.939

alignment_block:
US-09-680-121-2 x US-08-609-049A-27 ..
Align seg 1/1 to: US-08-609-049A-27 from: 1 to: 6831

53 PheValHisValLeuLysGlyValAlaAspIleTyrProGluAsnLeuAsnSe 69
|||||
5082 TTGTGTCGGTATATCGGGAATTACAGGAATTCATCAGAAAGATTATGCA 5131
|||||
69 rLysLysLysPheGly.....A 75
|||||
5132 TGCACCTTCCCTTGGTTAACTGCACAGTCTGCCGCTGCTGTCATGTG 5181
|||||
75 IaAspAspLysAsnGluValLysAsnLysProAlaValProLysAsnSer 91
|||||
5182 GCGCGTCCCATATCAATCCGTCGACGAGAAAGACGACTACTCTTATACA 5231
|||||
92 LeuHisLeuAspLeuGluLysArgAspLeuAsnGlyAsnPro...Ly 107
|||||
5232 GCGATTTTGAATATCGTTGTCATGCGTCGAGAGAAATAGCCATTCG 5281
|||||
107 sThrAsnLeuLysProGlySerProSerAspLeuGluAsnAla...ThrPro 123
|||||
5282 AGCTCGTTTACACATTTCTTTCACCCGCTGCTGCCGATCAGACAGAACCC 5331
|||||
124 LysLeuPheLeuGluGlyLysGlyLysSerValSerProGluSerLeu 140
|||||

```



```

205 ysVallysthrArgValleuArglysthrLeuaspProAlaheaspGlu 221
|||
5546 AACGCAAGACCAAGTGTGGCAGACCTGTGTGCCAGTTTCATGAGAA 5595
222 thrPheThrPheTyrGlylleProTyrThrGlnIleGlnGluLeuAlaLe 238
|||||
5596 ACTTTGAG...TACCGAATGCCACATGATATATATCAAGAGCGCCGCT 5642
238 uHisPheThrIleLeuSerPheaspArgPheSerArgaspAspIleLeu 255
|||||
5643 TCAGGTTCAGTTGGTGGCCACGACCCCTCGAGGAGAACGACGCTCTTG 5692
255 lylGluValleuIleProLeuSerGlylleGluLeuSerGluGlyMet 271
|||
5693 GA.....GGCTTCGATATGATCTGTGGAAGTAC 5721
272 LeuMetAsnArgGluIlellellysArgAsnValArglySerSerGlyAr 288
|||||
5722 GACCTGCGACAGAGCTCGTCGACCTGTATCGCCTGGCGGCTGTCCAG 5771
288 gGly.GluLeuLeuIleSerLeuGlyTyrGlnSerThrThrAsnThrLeu 304
|||||
5772 GAACCTACCGATCTGAGGACGAGCTATTGTGAACCTCTTGGACACTC 5821
305 ThrValValValleuLysAlaArgHisLeuPro 315
|||
5822 TGCCTACCGACATCAGCGCTAGGATATGCCA 5854

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-909-954-1

seq_documentation_block:
; Sequence 1, Application US/08909954A
; Patent No. 6100058
; GENERAL INFORMATION:
; APPLICANT: Allen, Maxine J.
; APPLICANT: Buckler, Alan J.
; TITLE OF INVENTION: GAP12 Genes and their Uses
; FILE REFERENCE: SEO-11P
; CURRENT APPLICATION NUMBER: US/08/909,954A
; CURRENT FILING DATE: 1997-08-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3148
; TYPE: DNA
; ORGANISM: H. sapiens
US-08-909-954-1

alignment_scores:
Quality: 136.50 Length: 194
Ratio: 1.437 Gaps: 8
Percent Similarity: 48.969 Percent Identity: 27.835

alignment_block:
US-09-680-121-2 x US-08-909-954-1 ..
Align seg 1/1 to: US-08-909-954-1 from: 1 to: 3148

154 LeuGlyThrLeuPheSerLeuGluTyrAsnPheGluArgLysAlaPh 170
|||||
285 CTAGGGCCCATGGCCAAAGACGACCTCCCTGAT..... 317
170 eValValAsnIleLysGluAlaArgGlyLeuProAlaMetAspGluGln 187
|||||
318 ....GTTCCGCTGTGTGAGCGCGCGCTGCCCAAGAC...GTGT 360
187 erMetThrSerAspProTyrIleLysMetThrIleLeuProGluLys 203
|||
361 CTGGGACGACGACCCCTACTGCTGTGAAAGTG.....GAGGACGAG 404
204 HisLysValLysThrArgValleuArglysthrLeuaspProAlaheas 220
|||||

```

```

405 GTGTGGCCGACGACACTACTGTGTGAGGAGCCGTGGCCCC..... 446
220 pGluThrPheThrPheTyrGlylleProTyrThrGlnIleGlnGluLeuA 237
|||||
447 .....TTCTGGGGGAGAGTACACGTCACCTGCTCTGG 483
237 lLeuHis.....PheThrIleLeuSerPheaspArgPheSerArg 250
|||||
484 ATTTCACACGACGTCGCCCTTCTACGTGCTGATGAGACACTGTCGGGAC 533
251 AspAspIlelleGlyGluValleuIleProLeuSerGlylleGluLeuSe 267
|||||
534 GACGACATCATC..... 545
267 rGluGlyLysMetLeuMetAsnArgGluIlellellysArgAsnValArg. 283
|||||
546 ....GGCAGATCTCGCTGAGGAGGAGGATTAACGCCGCCCGAG 591
284 .....LysSerSerGly 287
592 GGATTGACACTGATTACTGAGCCGAGTGCACCCAGATCGAAGTGG 641
288 ArgGlyGluLeuLeuIleSer.....LeuGlyTyrGlnSerThrThrAs 302
|||||
642 CAGGTGAGATCTGCTGCTGACGACAGATGCTGAGAGATGGGAGGCGG 691
302 nThrLeuThrValValleuLysAlaArgHisLeuProLysSerAspV 319
|||
692 CTGCTGCTGCCAGTGTCTTCAAGCCGACGAGCTGCTCCAGAGACA 741
319 aISerGlyLeuSerAspProTyrValLysVal 329
|||||
742 TCTCTGACACATCTGACCCATTGACAGCTGTG 773

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-539-205A-5

seq_documentation_block:
; Sequence 5, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3226 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear

```

MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 400..2901  
US-08-539-205A-5

alignment\_scores:  
Quality: 135.00 Length: 177  
Ratio: 1.392 Gaps: 6  
Percent Similarity: 54.802 Percent Identity: 25.424

alignment\_block:  
US-09-680-121-2 x US-08-539-205A-5 ..

Align seg 1/1 to: US-08-539-205A-5 from: 1 to: 3226

```
129 GUGUULYGLUSERVALSERPROGLUSERLEULYSERSETHSERIE 145
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
58 GGCAGGAGCCGGAGGGTCACCAAGTAGATTCCAGCAGCGCTAGT.. 105
145 UTHRSERGLUGLUGSLUGSLUGSLUGSLUGSLUGSLUGSLUGSLUG 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 ..CCAGCTGACACTTTCCAGCTGTTGTTTTCACG.... 138
162 LUTYASNPHEGLUARGLYSLALAPHEVALVALANILEYSLUALAARG 178
|||||:|||||:|||||:|||||:|||||:|||||:|||||
139 ..AGCTTTCAGGAAAG..... 153
179 GLYUAPROALAMELASPGLUGLINSERMETHRSERASPROTYRILEY 195
|||||:|||||:|||||:|||||:|||||:|||||:|||||
154 ..TATAGTATCGCTATGTGAA 173
195 SMETHRILEUPROGLULYSLSHLSYS.....VALYST 208
|||||:|||||:|||||:|||||:|||||:|||||:|||||
174 ACTTTCATTTAGTAGAGCGATGAGATACAGAACTTGCTTGGTCCAGA 223
208 HTRARGVALLEUARGLYSTRHLEUASPPROALAPHEASPLUTHRPHETHR 224
|||||:|||||:|||||:|||||:|||||:|||||:|||||
224 CAAAACATTTAAAGACACTGACCAACCAATGAGATGACAAATTTAT 273
225 PHEYRGLYILEPROTYRTHRGINILEGLINLEUALALENHIPHETH 241
|||||:|||||:|||||:|||||:|||||:|||||:|||||
274 TTCAGGCGTAACCCCTCTAATCAC.....AGACTCCTATTGA 311
241 RILEUSERPHEASPARPHEASERARGASPARLLEGLYGLUVAL 258
|||||:|||||:|||||:|||||:|||||:|||||:|||||
312 AGTATTTGACGAAATAGACTGACACGACGCGCTTCGTGGCCAGGTGG 361
258 EUILEUPROLEUSERGLYLEGLULEUSERGLULYSMET..... 271
|||||:|||||:|||||:|||||:|||||:|||||:|||||
362 AGCTGCCCTTAGTACCTTCGACAGAACATCCAAACATGGAGCGACCC 411
272 ..LEUMETASARGLUILEILEYARGASVALARGLYSSERSEGL 287
|||||:|||||:|||||:|||||:|||||:|||||:|||||
412 TATACATTAAAGACATTCTCCCTACACCAAGACATTAAGTCTCGAGT 461
287 VARGLYGLULEULEUSERLEUCYSTYR 297
|||||:|||||:|||||:|||||:|||||:|||||:|||||
462 TAAGGATTTTGGCATTTGAAATGGCCTAT 492
```

seq\_name: /cgn2\_6/prodata/2/lna/6A.COMB.seq:US-08-909-954-3

seq\_documentation\_block:

Sequence 3, Application US/08909954A  
Patent No. 6100058  
GENERAL INFORMATION:  
APPLICANT: Allen, Maxine J.  
APPLICANT: Buckler, Alan J.  
TITLE OF INVENTION: GAP2 Genes and their Uses  
FILE REFERENCE: SEO-11P  
CURRENT APPLICATION NUMBER: US/08/909,954A  
CURRENT FILING DATE: 1997-08-12

NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 3117  
TYPE: DNA  
ORGANISM: M. musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (0)...(0)  
US-08-909-954-3

alignment\_scores:  
Quality: 132.00 Length: 257  
Ratio: 1.091 Gaps: 13  
Percent Similarity: 47.082 Percent Identity: 24.514

alignment\_block:  
US-09-680-121-2 x US-08-909-954-3 ..

Align seg 1/1 to: US-08-909-954-3 from: 1 to: 3117

```
112 PROGLYSERPROSERASPLEUGLUSALATHRPROLYSLEUPHELEUG 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||
63 CCTCTGTGTCACGCTCAGAACCTCGGACGACTCTGCTGCTG..... 104
128 UGLYGLULYGLUSERVALSERPROGLUSERLEULYSERSETHSERL 145
|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 ..GAAAGTTGGAGCTGTGCGACATGCGCAGACGACGACGCGCTCGC 144
145 EUTHRSERGLUGLUGSLUGSLUGSLUGSLUGSLUGSLUGSLUGSLUG 161
|||||:|||||:|||||:|||||:|||||:|||||:|||||
145 TGACT..... 149
162 GLUTYASNPHEGLUARGLYSLALAPHEVALVALANILEYSLUALAAT 178
|||||:|||||:|||||:|||||:|||||:|||||:|||||
150 ..ATCCGCGTGTGGAGGAGG 169
178 GGLYUAPROALAMELASPGLUGLINSERMETHRSERASPROTYRILEY 195
|||||:|||||:|||||:|||||:|||||:|||||:|||||
170 AGCAGTCCCGCCAGAGAC..GTGCTGGAAGCAGTACCCCTATTGTC 216
195 YSMETHRILEUPROGLULYSLSHLSYSVALYSTHTRARGVALLEU 211
|||||:|||||:|||||:|||||:|||||:|||||:|||||
217 TGTGGAAGTG.....GATGACCAAGTGTGGCGAGACGACCAACCATC 260
212 ARGLYSTRHLEUASPPROALAPHEASPLUTHRPHETHRPHETRYGLY 228
|||||:|||||:|||||:|||||:|||||:|||||:|||||
261 TGGAGGAGCCTGAGCCCC.....TTTGGGGGGA 289
228 EPRTYRTHRGINILEGLINLEUALALEUHNIS.....PHETHR 242
|||||:|||||:|||||:|||||:|||||:|||||:|||||
290 GAGATACCCGCTTCCACCTTCATGTGACTCCACCCAGCTGAGCTAGC 339
242 LEUSERPHEASPARPHEASERARGASPARLLEGLYGLUVALLEU 258
|||||:|||||:|||||:|||||:|||||:|||||:|||||
340 TCGTGATGAGACACCTTGGACACGATGACATCATTTGGGAAGATGCA 389
259 ILEPROLEU.....SERGLYI 264
|||||:|||||:|||||:|||||:|||||:|||||:|||||
390 TTGAGCAAGAGCGATTCACAGCGACCCCTGAGGAGATCGACAGCTGANT 439
264 EGLULEUSER.....GLUGLYLSMETLEU 273
|||||:|||||:|||||:|||||:|||||:|||||:|||||
440 CAACTTGAGCCGAGTGATCCAGACGCTGAAGTACAGAGGTGAGTCTGCC 489
273 ELASNARGLUILEILEYARGASVALARGLYSSERSEGLYARGLY 289
|||||:|||||:|||||:|||||:|||||:|||||:|||||
490 TGGATGTGAAGCTATTGGAG.....GATGCTCGGGC 521
290 GLULEULEUSERLEUCYSTYRGLINSERTHRTHASNTHREUTHRYA 306
|||||:|||||:|||||:|||||:|||||:|||||:|||||
522 CGTGGCTC.....CGCTGCCAC..... 539
```

```

306 lvalleulysalarghisteuprolysseraspvalserglyleus 323
540 ...GAGACAGCCAGGAGCCTGGCCCCCGGAGCATCTCTGGACAT 585
323 easpProtyrVallysVal.....AsnLeutyRHIsalalys 336
586 CGGACCCATTGGCCCTGTCTTCTGGGGCAGC.....CATAGTTGGAA 629
337 ArgIleSerLysLysLysThr 343
630 ACTTCGACCATCAAGAACGAC 650

```

seq\_name: /cgn2\_6/ptodata/2/1na/6A.COMB.seq:US-08-895-601-2

```

seq_documentation_block:
; Sequence 2, Application US/08895601
; Patent No. 6060262
; GENERAL INFORMATION:
; APPLICANT: Beer-Romero, Peggy
; APPLICANT: Strack, Peter J.
; APPLICANT: Glass, Susan J.
; APPLICANT: Rolfe, Mark
; TITLE OF INVENTION: REGULATION OF KAPPA B (Kb) DEGRADATION,
; TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HONG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895.601
; FILING DATE: 16-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-096.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-7000
; TELEFAX: 617-832-1000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2790 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..2782
; US-08-895-601-2

```

```

alignment_scores:
Quality: 117.00      Length: 145
Ratio: 1.330        Gaps: 4
Percent Similarity: 60.690    Percent Identity: 22.759

```

alignment\_block:

```

US-09-680-121-2 x US-08-895-601-2
Align seg 1/1 to: US-08-895-601-2 from: 1 to: 2790

```

166 GluArgLysAlaPheValValAsnIleLysGluAlaArgLysLeuProAl 182

```

125 GAGAAATATTCACGAATGTGTGAGAGTAAGAGTTATAGCCGGAATAGGCCT 174
182 ametaSpGlnSerMet...ThSeraspProtyrIleLysMetThrI 198
175 TGCCAGAAAGCATATATTTGGAGCTAGTGAATCCTTAGAGAGTGCAGCT 224
198 leuLeuProGluLysLys.....HisLysValLysThrArgValLeu 211
225 TATATGACCCCAATGAATGAGGTTCTTACAGGTGTGCAAAACAAACCATTT 274
212 ArgLysThrLeuAspProAlaPheaspGluThrPheThrPheTyrglyI 228
275 AAAAAGAGTTTGAATTCAGAGTGAATGGAATGAAGAAATATATTATTCAGAGTCA 324
228 eProtyrThrGlnIleGlnIleuAlaLeuHisPheThrIleuSerP 245
325 TCCT.....CAGCAGCAGCGGCTTCTTTTGAAGGTGTTGACG 362
245 heaspArgPheSerArgAspAspIleIleGlyIleuAlaLeuIleProLeu 261
363 AAACCGATTGACAGAGAGATTTCTAGTCAAGTGAAGTTCACCTT 412
262 SerGlyIleGluLeuSerGlyLysMet.....LeuMetas 274
413 TATCCATTACCGACAGAAATCCAGATTTGAGAGACCATATACATTAA 462
274 nArgGluIleIleLysArgAsnValArgLysSerSerGlyArgGlyLul 291
463 GGAATTTGTTCTTCATCCAGAGAGTCAAAATCAAGAGTTAAAGTTATC 512
291 euLeuIleSerLeuCysTyrglnSerThrAsn 302
513 TGAGACTAAATGACTTATTATTACCTAAACCACT 547

```

seq\_name: /cgn2\_6/ptodata/2/1na/5B.COMB.seq:US-08-642-846-1

```

seq_documentation_block:
; Sequence 1, Application US/08642846
; Patent No. 5886151
; GENERAL INFORMATION:
; APPLICANT: HOSTETTER, MARGARET K.
; APPLICANT: GALE, CHERYL A.
; APPLICANT: BENDEL, CATHERINE M.
; APPLICANT: TAO, NIAN-JUN
; APPLICANT: KENDRICK, KATHLEEN
; TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
; TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUEITING, RAASCH, GERHARDT & SCHWAPPACH, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,846
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEITING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110.00280101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 1:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 5194 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-642-846-1

```

```

alignment_scores:
  quality: 102.50      length: 441
  ratio: 0.551         gaps: 24
  percent similarity: 42.177  percent identity: 20.181

```

```

alignment_block:
  US-09-680-121-2 x US-08-642-846-1 ..

```

```

Align seg 1/1 to: US-08-642-846-1 from: 1 to: 5194

```

```

40  ArgLysSerSerLysSerAsnLysThrPro.....Pr 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2443 CGTGTACTAGTGAATAAGTGAATCCCAATGCCATACATTCACAGAA 2492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50  ctyrltysphevalhlsvalleuylsglyvalasplletyrrprogluasnl 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2493 ATTCAAAGAGGAAATGATGATGACAGAGAGTTGTTAGTCCAGAC.... 2538
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
67  euasnserllysllyspheglyalaasplasplysasncluvallysasn 83
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2539 .....ATGATGATTTGATGATGATTCATCTCAATTT 2565
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
84  LysProalaValProLysAsnSerLeuHisLeuAspLeuGluLysArgAs 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2566 TTACCGAATATATCTGAAGACTCT.....GGATTAAAGA 2600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100  pleuasn...GlyAsnPherProLysThrAsnLeuLysProGlySer.... 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2601 TTTGAATTTTGCACAACTACCTCATAACACCAAGACCAAGACTTTTAA 2650
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
115  .....Proser 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2651 CTCCATTTGAGCACTAAATGCTCTTGCGAATATGATACAGATCTAAT 2700
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
117  AspleugluasnalatThrProLysLeuPheLeuGlu..... 128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2701 GTTGTGGAACCTCCTGAACCGAATATCATATGCTGAATATAGAAATGCTAG 2750
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129  .....GlyGluLysGluSerValSerP 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2751 ACGTTATACACTAATAAGCAGCGCCCAATTCAGCACCACCATTTGCCAC 2800
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
136  togLysSerLeuLysSerSerThrSerLeuThrSerGluGluLysGlnGlu 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2801 CACAAGCAGCAACATCTTCAACTGTTCCAAATTCAAATGAAGAGTGTCC 2850
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
153  LysLeuGlyThrLeuPhePheSerLeuGlu..... 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2851 ACATTTAGAGTGCCACATTTGAATTAGAAGAACTTCTTCAGCATTTAGC 2900
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
163  .....TyrAsnPherGluArgLysAlaPheValAlaAsnLeu 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2901 ACCTTGTGACATGTATATGATATTTTGTATGTTGCGTGGGGTGTCTA 2950
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
175  ys.....GluAlaArgGly.....LeuProLysMetLysP 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2951 AACCAATATTAAGGAGAGAAAGAAATGCAAGATGATGATGATGATGAT 3000
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
185  .....GluGlnSerMetThrSerAspPr 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3001 GATGATGTCAGAGAGATTTTGAATTCAGAAAGAGGTGATCTCAAGATGA 3050
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
192  ctyrltyleuylserthrleuProGlu.....LysLysHisLysValL 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

3051 ATATATTAATGCCAAACTTGTGATCAAAAAACCTAAAGAAATTCATTTG 3100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207  ysthrArgValLeuArgLysThrLeuAspProAlaPheAspGluThrPhe 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3101 TCACC.....GATCCCGAA...GACCGA..... 3120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224  ThrPheArgLysLeuProTyrThrGlnIleGlnGluLeuAla...LeuH 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3121 .....TATCAAGAAATTAACACAAACTGCTCTATAC 3152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239  spherThrIleLeuSerPheAspArgPheSerArgAspAspIleIleGlyG 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3153 CAATGCCACCATTTGATTCAGATTTATGCGCCGACAGACTCCATTTCTA 3202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
256  luValLeuIleProLeuSerGlyIleGlu..... 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3203 CCGACATGTGCTTATCTTATGATGATGAATGAAAAACACCATAGCGCT 3252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266  ..LeuSerGluLysMetLeuMetAsnArgGluLe.....IleLys 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3253 TTATATCTCTGATGCTGTTGTTTATGACACAGAGTACATCCGTTAA 3302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
279  sarGAsn...ValArgLysSerSerGlyArgGly.GluLeuLeuIleSer 294
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3303 ATCAAACTCTGTTTGTGTTCCACCGAGGGGACAGACAGCAATTAATCTT 3352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
295  LeuCySTyrGlnSer..... 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3353 CAATGTTACAGACCCAGATTTGATTAATCAATTCACCTGCTAGAAAT 3402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299  ..... 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3403 GTGCTGAACACAGTGAATGTCGCCATCAGTGAATGCTAGTACTAT 3452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
300  ....ThrThrAsnThrLeuThrValValLeuLysAlaArgHisLeu 314
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3453 TAGTTTAAACATTTGATGATGAATTTGATGACCAAGCTACAAATTTGTC 3502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
315  ProLysSerAspValSerGlyLeuSerAspProTyrVal..... 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3503 AAAAATTCACAGACACCTGCTTCAAAATCCGCCATFACTGTTCTGTGT 3552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
328  .....LysValAsnLeuThrHisAlaLysLysArgLies 339
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3553 GATGATGATGATGATGCGCCAGTGCACCTGAACAC.....CAAGAACTCC 3596
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
339  erLysLysLysThrHisValLysLys.....CysTh 349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3597 TACCAGAAAGAGGTCCATATTCAGACAGCCTGCGCAAGCTTCTTGTGCT 3646
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
349  pProAsnAlaValPheAsn 355
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
3647 CCCCTAGAAAATTCACCAAT 3665
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

seq_name: /cgn2_6/plodata/2/lna/SB_COMB.seq:US-08-392-625-16
seq_documentation_block:
; Sequence 16, Application US/08392625
; Patent No. 5837485

```

```

; GENERAL INFORMATION:
; APPLICANT: Entlan, Karl-Dieter
; APPLICANT: Gtz, Friedrich
; APPLICANT: Schnell, No. 5837485bert
; APPLICANT: Augustin, Johannes
; APPLICANT: Engelke, Gerhard
; APPLICANT: Rosenstein, Ralf
; APPLICANT: Kalesita, Cortina
; APPLICANT: Klein, Gora
; APPLICANT: Wleland, Bernd
; APPLICANT: Kupke, Thomas
; APPLICANT: Jung, G nther
; APPLICANT: Kellner, Roland
; TITLE OF INVENTION: Biosynthetic Process for The Preparation

```

```

: TITLE OF INVENTION: Of Chemical Compounds
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox
: STREET: 1100 New York Avenue
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/392,625
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/876,791
: FILING DATE: 30-APR-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmond, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 0652.0980002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8700 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-392-625-16

```

```

alignment_scores:
  Quality: 100.50      Length: 443
  Ratio: 0.493         Gaps: 22
  Percent Similarity: 46.050   Percent Identity: 20.767

```

alignment\_block:  
US-09-680-121-2 x US-08-392-625-16 ..

Align seg 1/1 to: US-08-392-625-16 from: 1 to: 8700

```

45 SerAsnLysThrProProlLysPheValHisValLeuLysGlyValAs 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2394 TCTATGAAATTACAAATAATTTCTAAATGCGCTTAGAAGAGAGATA 2443
61 pLleTyProGluAsnLeuAsnSerLysLysPheGlyAlaAspAspL 78
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2444 TAAATATTAGATTTGTGTAAT..... 2465
78 ysaAsnGluValLysAsnLysProlAlaValProlLysAsnSerLeuHisLeu 94
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2466 .....TTAATGTCACAAATTATTTGTCTTAAAACTATTTCCAATA 2507
95 AspLeuGluLysArgAspLeuAsnGlyAsnPheProLysThrAsnLeuL 111
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2508 GATTACCTATATA.....GATTATTCAGAGAAATGATTAATA 2542
111 sProGlySerProSerAspLeuGluAsnAlaThrProlLysLeuPheLeuG 128
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2543 ACAAGAGTTGAGCTGATATATTAGTGAAGCAGCATATATTCTCGTTA. 2591
128 LuGlyGluLysGluSerValSerProGluSerLeuLysSerSerThrSer 144
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2592 .....TTATCTCTATATCATTTTGGTACAAAACATATT 2624
145 LeuThrSerGluGluGluGluLysLeuGluLysThrLeuPhePheSerLe 161
:::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

```

```

2625 AGCAATTATCAGCAA.....TTTTTATGCA 2650
161 uGluTyAsnPheGluArgLysAlaPheValHisValLeuLysGlu... 176
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2651 TAAATATGATTTGACCAA.....CTAGTAATTTTAAAGCAATTCG 2691
177 .....AlaArgLysLeuProlAlaMetAspGluGlnSerMet 188
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2692 TCTCAGATATTAATGATTTGGCTATCCCAAAAAAGACAGTATAGTTT 2741
189 ThrSerAspProTyrlLysMetThrLLeuProGluLys..... 202
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2742 TCTAATAAC.....ATTGCATTTTAAAGAAAGATTGCT 2779
203 .....LysH 204
2780 TGCATTCACAAATACAGCCATATTGAATACAGAAACGACGTTAATA 2829
204 IsLysValLysThrArgValLeuArgLysThrLeuAspProlAlaPheAsp 220
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2830 ATTGAAGAAAGAAATATACAGTTCTAAATCATCGCCGCTTCAACT 2879
221 GluThrPheThr.....PheTyGlyLLeuProTyThrGlnLLeuGln 235
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2880 GAAATATATAGTGAATATATTGGAATTCATTAAGCGTTATGAGGA 2929
235 uLeuAlaLeuHisPheThrLLeu.....SerP 245
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2930 TTTGGCGTGATGATGATCCCATATTAGATCTTTAATGCCGCTCAACT 2979
245 heAspArgPheSerArgAspAspLLeuLLeuGluValLeuLLeuProLeu 261
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
2980 TTGAGAGTTTACGGGAAATTTCATATA..... 3008
262 SerGlyLLeuLeuSerGluGlyLysMetLeuMetAsnArgLLeuL 278
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3009 .....AAGAAAAAATCATATTCAAAAAAGAAATGT 3040
278 elys.....ArgAsnValArgLysSerSerGlyArgGlyLLeuLeuL 293
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3041 GCATCATTTACATTAATTACATGAATGAAATGCT.....TTAGAA 3081
293 LeSerLeuCySTyGlnSerThrThrAsnThrLeuThrValValValLeu 309
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3082 TAAGCAATTAAATGAAGAGCTCCTTAACGTCAGAAATGATATTTTG 3131
310 LysAlaArg.....HisLeuProlLysSe 317
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3132 AATATATATAGATATATATATCTGTTTAATTTAATTTACCTAATAAG 3181
317 AspValSerGlyLeuSerAspProTyVal.....LysV 329
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3182 TGATATAGAT...ATAATGACATATTTATTTGAGAGCTACATTTAAACAAC 3228
329 alaAsnLeuTyHisAlaLysLysArgLLeuSerLysLysThrHisVal 345
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3229 TTATATCTATAT.....TCTGAAGAACATGATTCACAGA 3260
346 LysLysCySThrProAsnAlaValPheAsn.....GluLe 357
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3261 ATTGATATCGTATCTAATTAATGATTTAATGATTTGATGATCGATTT 3310
357 uPheValPheAspLLeuProCyGluGlyLeuGluAspLLeuSerValGlu 373
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3311 ATACAAATTT.....TTAAGCAAAATTCATTATTGAAA 3342
374 .....PheLeuValLeuAspSerGluArgGly..... 382
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3343 AAACAAATTTATACACCTATTAACGAGAGAGGATGACTCATTTACT 3392
383 ..SerArgAsnGluValLLeuGlnLeuValLeuGlyAlaAlaAlaG 398
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
3393 TTTGTCCAGAAATTTATTAATAAATATTTTAAACACGCTACTTG 3442

```

Fri Dec 14 10:32:01 2001

us-09-680-121-2.p2n.rni

Page 14

```
398 uGlyThrGlyGlyGluHisTrpLysGluIleCysAspTyrProAlaArg 415
      :::::| | | | | :::::| | | | | :::::| | | | |
3443 GAAATTAATTTCAGAAATGTTTCTGAACTGAAATGTTAATAGGT 3492
      :::::| | | | | :::::| | | | | :::::| | | | |
415 InIleAla..... LysTrpHisVal 421
      :::::| | | | | :::::| | | | | :::::| | | | |
3493 TCGCAACTATTAGAGAAATGCAATATT 3521
```